

CR  
FE

RECEIVED

NOV -5 2003

# SEARCH REQUEST FORM

Access DB#

107753

Scientific and Technical Information Center

SEARCHER  
(STIC)

Requester's Full Name: Myron Hill Examiner #: 78884 Date: 11/4/03  
Art Unit: 1648 Phone Number 30 84521 Serial Number: 09/830 981  
Mail Box and Bldg/Room Location: 8E17 Results Format Preferred (circle): PAPER DISK E-MAIL  
8A16-office

If more than one search is submitted, please prioritize searches in order of need. MEJ

\*\*\*\*\*  
Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Cell Permeability Modulating Peptide

Inventors (please provide full names): 1) Hill, Myron  
2) Oess (Geb. Schmidt), Stephanie

Earliest Priority Filing Date: 11/3/98 Schmidt

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

C. Chan  
Rush

Please, 1) STN formula of claim 1. and

2) STN seq. of Fig. 1 which is: SEPIO<sup>#2</sup>

BEST AVAILABLE COPY

AA 12

SEQUENCE PL. ST. K. 11/11/03

## STAFF USE ONLY

### Type of Search

### Vendors and cost where applicable

Searcher: Bachly c4999 NA Sequence (#) STN  
Searcher Phone #:                      AA Sequence (#)                      Dialog                       
Searcher Location:                      Structure (#)                      Questel/Orbit                       
Date Searcher Picked Up:                      Bibliographic                      Dr. Link                       
Date Completed: 11-07-03 Litigation                      Lexis/Nexis                       
Searcher Prep & Review Time: 20 Fulltext                      Sequence Systems                       
Clerical Prep Time:                      Patent Family                      WWW/Internet                       
Online Time: 33 Other                      Other (specify) CGN

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 2003, 15:02:33 ; Search time 41 Seconds

(without alignments)  
46.457 Million cell updates/sec

Title: US-09-830-981-2

Perfect score: 61

Sequence: 1 PLSIFSRIGDP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03:\*

1: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA2002.DAT:\*

24: /SIDSI/gcgcdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1          | 61    | 100.0       | 12     | 21 | AA10598     |
| 2          | 61    | 100.0       | 12     | 21 | AA10598     |
| 3          | 61    | 100.0       | 12     | 23 | ABG78991    |
| 4          | 61    | 100.0       | 25     | 12 | AA14747     |
| 5          | 61    | 100.0       | 36     | 23 | AA17023     |
| 6          | 61    | 100.0       | 60     | 16 | AA176747    |
| 7          | 61    | 100.0       | 102    | 20 | AA130947    |
| 8          | 61    | 100.0       | 164    | 21 | AA13740     |
| 9          | 61    | 100.0       | 164    | 21 | AA190760    |

|    |    |       |     |    |          |
|----|----|-------|-----|----|----------|
| 10 | 61 | 100.0 | 174 | 12 | AA15617  |
| 11 | 61 | 100.0 | 174 | 13 | AA12367  |
| 12 | 61 | 100.0 | 215 | 21 | AA10597  |
| 13 | 61 | 100.0 | 281 | 13 | AA12471  |
| 14 | 61 | 100.0 | 281 | 20 | AA13283  |
| 15 | 61 | 100.0 | 389 | 13 | AA12472  |
| 16 | 61 | 100.0 | 389 | 13 | AA12474  |
| 17 | 61 | 100.0 | 393 | 22 | AA14097  |
| 18 | 59 | 96.7  | 55  | 8  | AA170002 |
| 19 | 59 | 96.7  | 64  | 22 | AA18133  |
| 20 | 59 | 96.7  | 64  | 23 | AA17027  |
| 21 | 59 | 96.7  | 67  | 22 | AA18135  |
| 22 | 59 | 96.7  | 67  | 23 | AA17029  |
| 23 | 59 | 96.7  | 174 | 12 | AA15618  |
| 24 | 59 | 96.7  | 174 | 13 | AA12386  |
| 25 | 59 | 96.7  | 174 | 24 | AA158074 |
| 26 | 59 | 96.7  | 347 | 21 | AA10596  |
| 27 | 59 | 96.7  | 389 | 22 | AA166931 |
| 28 | 59 | 96.7  | 389 | 24 | AA155073 |
| 29 | 54 | 88.5  | 174 | 12 | AA12394  |
| 30 | 54 | 88.5  | 174 | 12 | AA15622  |
| 31 | 54 | 88.5  | 174 | 13 | AA12871  |
| 32 | 54 | 88.5  | 174 | 24 | AA158070 |
| 33 | 54 | 88.5  | 174 | 24 | AA158071 |
| 34 | 54 | 88.5  | 174 | 24 | AA158072 |
| 35 | 54 | 88.5  | 174 | 24 | AA158073 |
| 36 | 54 | 88.5  | 281 | 7  | AA160163 |
| 37 | 54 | 88.5  | 281 | 7  | AA160560 |
| 38 | 54 | 88.5  | 281 | 7  | AA160617 |
| 39 | 54 | 88.5  | 281 | 8  | AA170294 |
| 40 | 54 | 88.5  | 281 | 15 | AA162870 |
| 41 | 54 | 88.5  | 281 | 20 | AA132834 |
| 42 | 54 | 88.5  | 281 | 22 | AA164419 |
| 43 | 54 | 88.5  | 389 | 7  | AA160794 |
| 44 | 54 | 88.5  | 400 | 17 | AA193801 |
| 45 | 54 | 88.5  | 400 | 21 | AA154045 |

## ALIGNMENTS

|          |   |                           |
|----------|---|---------------------------|
| RESULT 1 | AA10598   | AA10598                   |
| ID       | AA10598   | standard; Protein; 12 AA. |
| AC       | AA10598;  |                           |
| XX       |   |                           |
| DT       | 08-JAN-2001   | (first entry)             |
| DE       | HBV fusion protein cell permeability peptide fragment.              |                           |
| XX       |   |                           |
| KW       | Fusion protein; protein coat; virus-specific packaging signal; psi; |                           |
| KW       | virus protein; cell permeability; cell-specific binding site; LMB;  |                           |
| KW       | large surface protein; core antigen; gene therapy.                  |                           |
| XX       |   |                           |
| OS       | Hepatitis b virus.  |                           |
| XX       |   |                           |
| PN       | WO200046376-A2.   |                           |
| XX       |   |                           |
| PD       | 10-AUG-2000.  |                           |
| XX       |   |                           |
| PF       | 04-FEB-2000; 2000MO-DE00363.  |                           |
| XX       |   |                           |
| PR       | 05-FEB-1999; 99DE-1004800.  |                           |
| XX       |   |                           |
| PA       | (HILD/) HILD E.   |                           |
| XX       |   |                           |
| PI       | Hildt E, Hofecheider P;   |                           |
| XX       |   |                           |
| DR       | WPI; 2000-514959/46.  |                           |
| XX       |   |                           |
| XX       | N-PSDB; AA17151.  |                           |
| PT       | Particle for cell-specific gene delivery, useful in gene therapy.   |                           |

PT comprises nucleic acid in protein coat that includes a fusion protein  
 of viral protein, permeability peptide and cell-binding site -  
 PS Disclosure, Page 2, 34pp; German.

CC This invention describes a novel particle (A), comprising a protein coat  
 CC with a fusion protein (FP), and, inside the coat, a nucleic acid (I)  
 CC including the sequence for a virus-specific packaging signal (psi) and a  
 CC structural gene. FP contains a virus protein (VP), a peptide (P) that  
 CC mediates cell permeability and a heterologous cell-specific binding site  
 CC (RGD). The invention also describes (1) producing (A) in which FP  
 CC contains an LHBs (large surface protein of hepatitis B virus (HBV)) and  
 CC RGD; (2) preparing (A) in which FP contains an HBV core antigen (HBcAg),  
 CC (P) and RGD; (3) FP; (4) DNA encoding FP; and (5) expression vector  
 CC containing the DNA of (d). The products of the invention are used in gene  
 CC therapy of cells and tissues, in vivo or ex vivo. This sequence  
 CC represents a fragment of the Hepatitis B virus cell permeability peptide  
 CC which is described in the method of the invention.

CC Sequence 12 AA;

Query Match 100.0%; Score 61; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSTFSRIGDP 12  
 |||||  
 1 PLSSTFSRIGDP 12

DB

RESULT 2

AA87906  
 ID AAY87906 standard; Protein; 12 AA.

AC AAY87906;

DT 01-SEP-2000 (first entry)

DE HBV cell surface protein ZPP fragment.

KW Cell surface protein; ZPP; cell permeability mediating polypeptide;  
 therapy; treatment.

OS Hepatitis B virus.

PN DE19850718-C1.

PD 18-MAY-2000.

PF 03-NOV-1998; 98DE-1050718.

PR 03-NOV-1998; 98DE-1050718.

PA (HILD/) HILD E.

PI Hildt E, Oess S;

DR WPI, 2000-340689/30.

DR N-PSDB; AAA95621.

PT Novel cell permeability-mediating polypeptide useful for mediating  
 PT permeability of substances such as other polypeptides in cells -

PS Claim 1, Fig 1, 8pp; German.

CC This invention describes a novel cell permeability mediating polypeptide  
 CC (CPM) comprising (A) or a sequence which differs by one or more amino  
 CC acids, which shows cell permeability and is not hepatitis B virus (HBV)  
 CC surface protein and where the DNA sequence of the latter amino acid  
 CC sequence hybridizes with the DNA of (I). CPM can be used to mediate cell  
 CC permeability of substances, especially polypeptides, nucleic acids and  
 CC chemical connections. This is useful for increasing the reactivity of  
 CC such substances in therapeutic treatment of various conditions. This

CC sequence represents a fragment of the HBV cell surface protein ZPP which  
 CC is described in the method of the invention.

CC Sequence 12 AA;

Query Match 100.0%; Score 61; DB 21; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSTFSRIGDP 12  
 |||||  
 1 PLSSTFSRIGDP 12

DB

RESULT 3

ABG78991  
 ID ABG78991 standard; Peptide; 12 AA.

AC ABG78991;

DT 15-NOV-2002 (first entry)

DE Cell penetrating peptide PRES.

KW Cell penetrating peptide; cancer; tumour; melanoma; thymoma;  
 KW Lymphoma; sarcoma; lung cancer; non-Hodgkin's lymphoma; leukemia;  
 KW Hodgkin's lymphoma; uterine cancer; cervical cancer; bladder cancer;  
 KW kidney cancer; adenocarcinoma; breast cancer; prostate cancer;  
 KW ovarian cancer; pancreatic cancer; vaccine; dendritic cell;  
 KW tumour infiltrating lymphocyte; TIL; human leukocyte antigen; HLA;  
 KW cytostatic.

OS Hepatitis B virus.

PN WO200264057-A2.

PD 22-AUG-2002.

PF 15-FEB-2002; 2002WO-US05212.

PR 15-FEB-2001; 2001US-268687P.

PA (BAYU ) BAYLOR COLLEGE MEDICINE.

PI Wang R;

DR WPI, 2002-627577/67.

PT Novel composition for treating a disease in an animal, comprises an  
 PT immune effector cell and cell penetrating peptide associated with an  
 PT antigen or antibody -

PS Disclosure, Page 11, 61pp; English.

CC The invention relates to a composition (I) comprising an immune effector  
 CC cell and a cell penetrating peptide (CPP) associated with an antigen or  
 CC antibody. Also included are (1) a vaccine comprising (I), CPP  
 CC associated with an antigen, and a pharmaceutically acceptable carrier  
 CC and (2) preparing a composition for a disease, by providing (I)  
 CC and CPP associated with an antigen for disease, and introducing the  
 CC antigen-associated CPP to (I), where antigen enters into the cell.  
 CC The antigens are, for example, tumour antigen derived epitopes  
 CC recognised by tumour infiltrating lymphocytes (TIL) of HLA (human  
 CC leukocyte antigen) class I or II. The composition is useful for enhancing  
 CC immunity in an animal to a disease, by administering a mature dendritic  
 CC cell comprising CPP associated with an antigen to disease, to the animal,  
 CC such that following the administration, animal is protected from disease,  
 CC where the animal comprises both CD4+ and CD8+ T cells. It is also useful  
 CC for treating a disease (e.g. cancer, tumour, melanoma, thymoma,  
 CC lymphoma, sarcoma, lung cancer, non-Hodgkin's lymphoma, leukemia,  
 CC Hodgkin's lymphoma, uterine cancer, cervical cancer, bladder cancer,  
 CC kidney cancer, adenocarcinoma, breast cancer, prostate cancer,  
 CC ovarian cancer and pancreatic cancer). The animal is further subjected to

CC a cancer treatment including surgery, radiation, chemotherapy or gene  
CC therapy. The administration of (1), preferably dendritic cell is prior  
CC to, subsequent to or concurrent with, the cancer treatment. The present  
CC sequence is cell penetrating peptide of the invention.

XX Sequence 12 AA;

Query Match 100.0%; Score 61; DB 23; Length 12;  
Best Local Similarity 100.0%; Pred. No. 4.5e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSRIGDP 12  
|||  
Db 1 PLSISFSRIGDP 12

RESULT 4

AAR14747  
ID AAR14747 standard; peptide; 25 AA.

XX AAR14747;

AC 28-JAN-1992 (first entry)

XX HBV Pre S2 peptide.

DE Monoclonal antibodies; hepatitis B virus; envelope; vaccine;

XX epitope; antigen.

OS Synthetic.

XX EP456215-A.

XX 13-NOV-1991.

PD 08-MAY-1991; 91EP-0107488.

XX 11-MAY-1990; 90US-0522286.

XX (ABBO) ABBOTT LABORATORIES.

PI Mims LT, Floreani MF;

XX WPI; 1991-334198/46.

DR New MAb to Pres2 and Pres1 polypeptide(s) of Hepatitis-B

XX envelope - used to sub-type HBV in sample and as vaccine against

PT HBV.

PS Claim 1, Page 22; 24pp; English.

CC The peptide corresponds to residues 150-174 of the HBV Pres2 protein.

CC It was used to raise monoclonal antibodies specific for the HBV M

CC protein but which do not bind to the HBV S or L proteins. The MAb

CC can be used to sub-type HBV and to prepare subunit vaccines for HBV.

XX Sequence 25 AA;

Query Match 100.0%; Score 61; DB 12; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.0001;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSRIGDP 12  
|||  
Db 11 PLSISFSRIGDP 22

RESULT 5  
AAE17023  
ID AAE17023 standard; peptide; 36 AA.  
XX AAE17023;  
XX

DT 18-APR-2002 (first entry)

XX Hepatitis B virus (HBV) ayw subtype pre-S2 region.

DE Hepatitis B virus; HBV, core antigen, HBcAg; immune system; typhoid;

XX Hepatitis B virus; HBV, core antigen, HBcAg; immune system; typhoid;

KW prophylactic; gene therapy; vaccine; hepatitis A virus; HAV; herpes;

KW hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;

KW tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;

KW dengue fever; yellow fever; malaria; whooping cough; salmonellosis;

KW food poisoning; meningitis; gonorrhea; antiviral; antibacterial;

KW antiprotocozal; pre-S2 region.

XX Hepatitis B virus.

XX WO200198333-A2.

XX 27-DEC-2001.

XX 22-JUN-2001; 2001WO-GB02817.

XX 22-JUN-2000; 2000GB-0015308.

XX 06-OCT-2000; 2000GB-0024544.

XX (CELL-) CELTECH PHARM LTD.

XX Page M, Li J, Pumpens P;

XX WPI: 2002-098223/13.

XX N-PSDB; AMD27424.

XX New proteins comprising a modified hepatitis B core antigen, useful as

XX a vaccine in prophylactic or therapeutic vaccination of the human or

XX animal body, particularly against hepatitis B virus infection -

XX Disclosure; Fig 3A; 40pp; English.

XX The invention relates to modified proteins comprising hepatitis B virus

XX (HBV) core antigen (HBcAg) wherein one or more of the four arginine

XX repeats has been deleted and the protein comprising the C-terminal

XX cysteine of HBcAg. The deleted region may be replaced by an epitope

XX from a protein other than HBcAg, in which case the HBcAg acts as a

XX carrier to present the epitope to the immune system. This chimeric

XX protein or its nucleic acid is useful as a vaccine or in a method of

XX prophylactic or therapeutic vaccination of the human or animal body,

XX particularly against HBV. The nucleic acid encoding the protein may

XX be used in gene therapy or DNA vaccination protocols. The chimeric

XX protein or its nucleic acid may also be used as the basis of a

XX prophylactic vaccine against a range of diseases, e.g. HBV, hepatitis

XX A virus (HAV), hepatitis C virus (HCV), influenza, foot-and-mouth

XX disease, polio, herpes, rabies, acquired immunodeficiency syndrome

XX (AIDS), dengue fever, yellow fever, malaria, tuberculosis, whooping

XX cough, salmonellosis, typhoid, food poisoning, diarrhoea, meningitis

XX or gonorrhea. The present sequence is Hepatitis B virus (HBV) ayw

XX subtype pre-S2 region.

XX Sequence 36 AA;

Query Match 100.0%; Score 61; DB 23; Length 36;  
Best Local Similarity 100.0%; Pred. No. 0.00015;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSRIGDP 12  
|||  
Db 22 PLSISFSRIGDP 33

RESULT 6  
AAR76747  
ID AAR76747 standard; peptide; 60 AA.  
XX AAR76747;  
XX 18-MAR-1996 (first entry)  
DT

```

XX DE HBV surface antigen pre-S2 region and FimH residues 224-226.
XX
XX FimH; type 1 fimbriae; organelle; adhesin; alpha-D-mannoside residue;
XX FimA; FimF; FimG; receptor binding site; PCR; amplicity.
XX OS Chimeric - Hepatitis B virus.
XX OS Chimeric - Escherichia coli.
XX
XX Key Location/Qualifiers
XX F1 Misc-difference 1 /note= "Represents FimH residue 224"
XX F1 Peptide 2..4 /note= "Linker peptide"
XX F1 Peptide 5..56 /note= "HBV surface antigen pre-S2 region"
XX F1 Peptide 57..59 /note= "Linker peptide"
XX F1 Misc-difference 60 /note= "Represents FimH residue 226"
XX
XX PN W09520657-A1.
XX PD 03-AUG-1995.
XX
XX PF 27-JAN-1995; 95WO-DK00042.
XX
XX PR 27-JAN-1994; 94US-0187166.
XX
XX PA (GXBI-) GX BIOSYSTEMS AS.
XX
XX PI Hasty DL, Klemm P, Molin S, Pallesen L, Sokurenko EV;
XX
XX DR WPI: 1995-275442/36.
XX
XX DR N-PSDB; AAQ93059.
XX
XX PT Receptor specific bacterial adhesins - useful for targeting active
XX PT compounds and microbial cells to locations of receptors
XX
XX PS Example 3; Page 57; 152pp; English.
XX
XX CC This sequence is encoded by a fragment of the plasmid pLPA38, which
XX CC comprises a fimb fusion gene comprising the sequence encoding the pre-S2
XX CC region of the hepatitis B viral surface antigen inserted into the fimb
XX CC gene. The primer sequences given in AAQ93057-58 caused the insertion of
XX CC the pre-S2 region into the FimH protein at position 224-226. The
XX CC resulting chimeric gene was then further modified by insertion of the
XX CC cholera toxin B chain into a different position of the FimH adhesin of
XX CC type 1 fimbriae. Restriction site handles (BglII-sites) were introduced
XX CC into the fimb gene, and the foreign epitopes are then inserted in-frame.
XX CC In the selected positions the insertion of the epitopes did not
XX CC significantly alter the adhesive function of the FimH protein. The
XX CC expression of the chimeric proteins on the surface of fimbriae on
XX CC bacterial hosts illustrated the possibility of using bacterial adhesins
XX CC as general presenters of foreign antigens and epitopes. These chimeric
XX CC genes may be used in the production of variant FimH adhesins which may
XX CC be useful for targeting active compounds and microbial cells to
XX CC locations comprising selected receptors to which the adhesins bind.
XX
XX SQ Sequence 60 AA;
XX
XX Query Match 100.0%; Score 61; DB 16; Length 60;
XX Best Local Similarity 100.0%; Pred. No. 0.00026;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 PLSSIFSRIGDP 12
XX |||||||||||
XX Db 42 PLSSIFSRIGDP 53
XX
XX RESULT 7
XX AA30947
XX ID AAY30947 standard; Protein; 102 AA.

```

```

XX AC AAY30947;
XX
XX DT 20-OCT-1999 (first entry)
XX
XX DE Human HbsAg (PreS2 plus S region) truncated peptide.
XX
XX KW Vaccine; accessory molecule; antigen; transport; presentation; allergy;
XX KW antibacterial; antiviral; antifungal; anti-allergic; antidiabetic;
XX KW anti-inflammatory; anti-arthritic; anti-asthma; anticancer; treatment;
XX KW prevention; infection; bacterial; viral; fungal; autoimmune disease;
XX KW rheumatoid arthritis; diabetes; multiple sclerosis; pancreatitis;
XX KW inflammatory condition; psoriasis; immune deficiency; metastases;
XX KW cancer; Th1-type helper response; humoral; cellular response; B cell;
XX KW T cell; proliferation; immunoglobulin synthesis; isotype switching;
XX KW immunomodulator; immune response; asthma; human; HbsAg.
XX
XX OS Homo sapiens.
XX
XX PN W09941368-A2.
XX
XX PD 19-AUG-1999.
XX
XX PF 10-FEB-1999; 99WO-US03020.
XX
XX PR 11-FEB-1998; 98US-0074294.
XX
XX PR 11-FEB-1998; 98US-0021769.
XX
XX PA (MAXY-) MAXYGEN INC.
XX
XX PI Howard R, Punnonen J, Stemmer WPC, Whalen RG;
XX
XX DR WPI: 1999-508645/42.
XX
XX PT Identifying nucleic acid that directly or indirectly modulates the
XX PT immune response to a genetic vaccine vector, e.g. for prevention of
XX PT infection or cancer
XX
XX PS Example 3; Page 70; 105pp; English.
XX
XX CC This invention describes the identification of a novel polynucleotide (I)
XX CC that modulates the immune response to a genetic vaccine vector (A), or
XX CC encodes a polypeptide (II) with similar effect. The invention also
XX CC describes the identification of a polynucleotide (Ia) encoding an
XX CC accessory molecule (Iia) that improves transport and presentation of
XX CC antigen by a cell. The products of the invention have antibacterial,
XX CC antiviral, antifungal, anti-allergic, antidiabetic, anti-inflammatory,
XX CC anti-arthritic, anti-asthma, anticancer and immunomodulatory. Optimized
XX CC (I) are incorporated into (A), or (I) or its encoded (II) are
XX CC administered together with (A). (A) are used to treat or prevent
XX CC infections (bacterial, viral or fungal), autoimmune disease (e.g.
XX CC rheumatoid arthritis, diabetes or multiple sclerosis), other inflammatory
XX CC conditions (e.g. psoriasis or pancreatitis), immune deficiency, allergy,
XX CC asthma or cancer (including metastases). (I) are also used for
XX CC recombinant production of (II). (I) make it possible to tailor an immune
XX CC response to particular requirements, e.g. to direct a Th1-type helper
XX CC response to increase humoral or cellular responses (functioning as
XX CC adjuvant) to control B or T cell proliferation to induce immunoglobulin
XX CC synthesis or isotype switching. This sequence represents the HbsAg
XX CC (PreS2 plus S region) truncated peptide which is used in the method of
XX CC the invention.
XX
XX SQ Sequence 102 AA;
XX
XX Query Match 100.0%; Score 61; DB 20; Length 102;
XX Best Local Similarity 100.0%; Pred. No. 0.00046;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 PLSSIFSRIGDP 12
XX |||||||||||
XX Db 41 PLSSIFSRIGDP 52
XX
XX RESULT 7
XX AA30947
XX ID AAY30947 standard; Protein; 102 AA.

```

```

RESULT 8
AA93740
ID AAY93740 standard; Protein; 164 AA.
XX
XX AAY93740;
AC
XX 03-OCT-2000 (first entry)
DT
XX
XX Amino acid sequence for pre-S peptide.
DE
XX Plant expression vector; antigen; hepatitis B surface antigen; HBsAg;
XX vaccine; ss.
XX
XX Synthetic.
OS
XX WO200037610-A2.
XX
XX 29-JUN-2000.
PD
XX
XX 23-DEC-1999; 99WO-US31020.
PF
XX
XX 23-DEC-1998; 98US-0113827.
PR
XX (BOYC-) BOYCE THOMPSON INST PLANT RES.
PA
XX (HEAL-) HEALTH RES INST.
XX
XX Mason HS, Thanavala Y, Arentzen CJ, Richter E;
PI
XX WPI; 2000-452181/39.
DR
XX N-PSDB; AAA47000.
XX
XX New expression vector for transforming plants comprising two expression
PT
XX cassettes useful for producing plant material comprising anti-hepatitis
PT
XX B antibodies -
XX
XX Disclosure; Fig 5; 144p; English.
XX
XX PS
XX
XX The specification describes a plant expression vector which comprises
CC
XX two expression cassettes, the first comprising a polynucleotide encoding
CC
XX an antigen and the second comprising a non-identical polynucleotide
CC
XX encoding the same antigen. The antigen is especially a hepatitis B
CC
XX surface antigen (HBsAg). The expression vector is used to transform
CC
XX bacterial and plant cells to elicit the production of anti-hepatitis B
CC
XX antibodies and are useful as vaccines. The present sequence represents
CC
XX a pre-S peptide.
XX
XX SQ
XX
XX Sequence 164 AA;
XX
XX
XX Query Match
XX Best Local Similarity 100.0%; Score 61; DB 21; Length 164;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY
XX 1 PLSISFSRIGDP 12
XX |||||
XX DB 149 PLSISFSRIGDP 160
XX
XX
XX RESULT 9
XX AAY90760
XX ID AAY90760 standard; Protein; 164 AA.
XX
XX AAY90760;
AC
XX
XX 18-AUG-2000 (first entry)
DT
XX
XX Hepatitis B virus pres protein sequence.
DE
XX
XX Nuclear localisation signal; NLS; SV40; large T cell antigen;
XX fusion protein; diagnosis; detection; variant; vaccine; chaperone;
XX coprecipitate.
XX
XX KM
XX Hepatitis B virus.
XX
XX OS
XX

```

```

PN WO200020606-A1.
XX
XX 13-APR-2000.
PD
XX
XX 02-OCT-1998; 98WO-EP06298.
PF
XX
XX 02-OCT-1998; 98WO-EP06298.
PR
XX
XX (REIM/) REIMANN H.
PA
XX (SCHI/) SCHIRMECK R.
XX
XX Reimann H, Schirmeck R;
PI
XX
XX WPI; 2000-328936/28.
DR
XX
XX Novel polynucleotide encoding a fusion protein which is stable in a
PT
XX cell, useful for the production of peptides which coprecipitate
PT
XX chaperones, by using truncated variants of the SV40 large T antigen
PT
XX with an intact N-terminus -
XX
XX PS
XX Example 1; Page 26; 57pp; English.
XX
XX CC
XX The present invention describes a polynucleotide (I) encoding a fusion
CC
XX protein (Pp) (II) which is stable in a cell, and comprises a first
CC
XX polypeptide (P1) and a second polypeptide (P2) which co-precipitates a
CC
XX chaperone. The invention also provides methods for the production of
CC
XX fusion proteins which comprise epitopes. Compositions comprising the
CC
XX fusion proteins or epitopes of the invention are used as vaccines, and
CC
XX for the production of antibodies. The methods provide an inexpensive
CC
XX and efficient means of reliably expressing (poly)peptides comprising
CC
XX epitopes. The present sequence represents the Hepatitis B virus (HBV)
CC
XX pres protein sequence, which is used in an example from the present
CC
XX invention.
XX
XX SQ
XX
XX Sequence 164 AA;
XX
XX
XX Query Match
XX Best Local Similarity 100.0%; Score 61; DB 21; Length 164;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY
XX 1 PLSISFSRIGDP 12
XX |||||
XX DB 150 PLSISFSRIGDP 161
XX
XX
XX RESULT 10
XX AAR15617
XX ID AAR15617 standard; Protein; 174 AA.
XX
XX AAR15617;
AC
XX
XX 25-MAR-2003 (updated)
DT
XX 25-MAR-1992 (first entry)
DT
XX
XX HBsAg pre-S region subtype ayw.
DE
XX
XX T-cell epitope; vaccine; hepatitis B virus; antigen.
XX
XX OS
XX Synthetic.
XX
XX WO9117768-A.
XX
XX 28-NOV-1991.
PD
XX
XX 10-MAY-1991; 91WO-US03268.
PF
XX
XX 11-MAY-1990; 90US-0522663.
PR
XX
XX (SCHI-) SCHIRPS CLINIC & RE.
PA
XX
XX Milich DR, Thornton GB;
PI
XX
XX WPI; 1991-369007/50.
XX

```

XX Hepatitis B virus surface antigen epitope(s) - useful as vaccines,  
PT immunogens or diagnostic reagents  
XX  
PS Claim 1, Fig. 1, 91pp; English.  
XX  
CC The amino acid sequence is that of a pre-S T cell epitope polypeptide  
CC of the pre-S (2) region of hepatitis B surface antigen Y (HBsAg/Y).  
CC It can be used to prime or vaccinate a host to induce responsiveness  
CC to HBV vaccine. The T cell epitope polypeptides can also be used as  
CC immunogens that prime T cells that respond to native HBsAg B cell  
CC epitope polypeptide. The T cell epitope polypeptides are also useful  
CC as substitutes for carrier immunogens such as KLH and are safe,  
CC defined and T cell-active. In addition to their use as vaccines, the  
CC polypeptides can be used as immunogens for prodn. of antibodies. See  
CC also AAR15618-R15622.  
CC (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 174 AA;  
XX  
Query Match 100.0%; Score 61; DB 12; Length 174;  
Best Local Similarity 100.0%; Pred. No. 0.00082;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 PLSISFRRIGDP 12  
DB 160 PLSISFRRIGDP 171  
XX  
RESULT 11  
ID AAR23867 standard; protein; 174 AA.  
XX  
AC AAR23867;  
XX  
DT 25-MAR-2003 (updated)  
DT 11-NOV-1992 (first entry)  
XX  
DE Pre-S gene region translation product (1).  
XX  
KM env gene; hepatitis B virus; HBV; S-protein; hydrophilic;  
KM hepanavirus; immune response; antibody; envelope.  
XX  
OS Hepatitis B virus, subtype ayw.  
XX  
PN EP485361-A.  
XX  
PD 13-MAY-1992.  
XX  
PF 25-APR-1987; 87BP-0100663.  
XX  
PR 28-APR-1986; 86US-0856522.  
XX  
PA (NYBL-) NY BLOOD CENT INC.  
PA (CALY) CALIFORNIA INST OF TECHN.  
XX  
PI Kent SBH, Neurath AR;  
XX  
DR WPI; 1992-161100/20.  
XX  
PT Hepatitis B pre-S peptide immunogen and vaccine - for treatment  
PT and diagnosis of hepatitis B  
XX  
PS Disclosure; Fig 2; 59pp; English.  
XX  
CC The sequences given in AAR23867 - AAR23871 are amino acid sequences  
CC deduced from sequences of the pre-S portion of the env genes  
CC corresponding to several hepatitis B virus (HBV) subtypes. These  
CC proteins have properties distinct from those of the S-protein in  
CC that they have high hydrophilicity and a high percentage of charged  
CC residues, an absence of Cys residues, the highest subtype-dependant  
CC variability among HBV DNA gene products and little homology with  
CC analogous sequences corresponding to nonhuman hepadnavirus. These

CC properties suggest that the pre-S gene coded portion of the HBV  
CC envelope is expose to the surface of the virion, is a target for  
CC the hosts immune response and is responsible for the host range of HBV  
CC (limited to humans and some primates). Synthetic peptides and  
CC antibodies against them, having predetermined specificity offer the  
CC opportunity to explore the biological role of the pre-S protein moiety  
CC of the HBV envelope. Portions of these amino acid sequences can be  
CC used in a vaccine or in diagnostics for the detection of antigens and  
CC antibodies, esp. those for the pre-S gene in sera of HBV infected  
CC humans and certain animals, eg. chimpanzees.  
CC (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 174 AA;  
XX  
Query Match 100.0%; Score 61; DB 13; Length 174;  
Best Local Similarity 100.0%; Pred. No. 0.00082;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 PLSISFRRIGDP 12  
DB 160 PLSISFRRIGDP 171  
XX  
RESULT 12  
ID AAB10597 standard; protein; 215 AA.  
XX  
AC AAB10597;  
XX  
DT 08-JAN-2001 (first entry)  
XX  
DE HBV fusion protein comprising HBcAg and RGD.  
XX  
KM Fusion protein; protein coat; virus-specific packaging signal; psi;  
KM virus protein; cell permeability; cell-specific binding site; LHB;  
KM large surface protein; core antigen; gene therapy.  
XX  
OS Hepatitis b virus.  
XX  
PN WO200046376-A2.  
XX  
PD 10-AUG-2000.  
XX  
PF 04-FEB-2000; 2000WO-DE00363.  
XX  
PR 05-FEB-1999; 99DE-1004800.  
XX  
PA (HILD/) HILDT E.  
XX  
PI Hildt E, Hofschneider P;  
XX  
DR WPI; 2000-514959/46.  
DR N-PSDB; AAA71735.  
XX  
PT Particle for cell-specific gene delivery, useful in gene therapy,  
PT comprises nucleic acid in protein coat that includes a fusion protein  
PT of viral protein, permeability peptide and cell-binding site  
XX  
PS Claim 14; Fig 2; 34pp; German.  
XX  
CC This invention describes a novel particle (A), comprising a protein coat  
CC with a fusion protein (FP), and, inside the coat, a nucleic acid (1)  
CC including the sequence for a virus-specific packaging signal (psi) and a  
CC structural gene. FP contains a virus protein (VP); a peptide (P) that  
CC mediates cell permeability and a heterologous cell-specific binding site  
CC (RGD). The invention also describes (1) producing (A) in which FP  
CC contains an LHBs (large surface protein of hepatitis B virus (HBV)) and  
CC RGD; (2) preparing (A) in which FP contains an HBV core antigen (HBcAg),  
CC (P) and RGD; (3) FP; (4) DNA encoding FP; and (5) expression vector  
CC containing the DNA of (d). The products of the invention are used in gene  
CC therapy of cells and tissues, in vivo or ex vivo. This sequence  
CC represents a fusion protein which is described in the method of the



```

CC invention.
XX
SQ Sequence 215 AA;
Query Match 100.0%; Score 61; DB 21; Length 215;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
   |||||
DB 2 PLSIFSRIGDP 13

RESULT 13
AAR27471
ID AAR27471 standard; Protein; 281 AA.
XX
AC AAR27471;
XX
DT 25-MAR-2003 (updated)
DT 24-FEB-1993 (first entry)
XX
DE spsAg protein.
XX
KW Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;
KW vaccinia virus; H6; early/late; promoter; NYVAC; recombinant; HBV L;
KW large pre-S antigen; lpsAg; fusion protein; pre-S region; S12/core;
KW S1; S2; Copenhagen vaccine strain; vaccinia virus; virulence factor;
KW deletion loc1; recipient loc1.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..55
FT /label= S2
FT Region 56..281
FT /label= S
XX
PN MO9215672-A1.
XX
PD 17-SEP-1992.
XX
PF 09-MAR-1992; 92WO-US01906.
XX
PR 07-MAR-1991; 91US-0666056.
PR 11-JUN-1991; 91US-0713867.
PR 06-MAR-1992; 92US-0847951.
XX
PA (VIRO-) VIROGENETICS CORP.
XX
PI Cox WI, De Taisne C, Francis J, Gettig RR, Johnson GP,
PI Limbach KJ, Norton EK, Paciotti E, Perkins ME, Pincus SE,
PI Riviere M, Tartaglia J, Taylor J;
XX
DR WPI; 1992-331718/40.
DR N-PSDB; AAQ29103.
XX
PT Vaccine comprises recombinant, attenuated pox-virus - use for
PT vaccinating against viral infections such as rabies, hepatitis B,
PT HIV, HSV, EBV, CMV, mumps etc.
XX
PS Disclosure; Fig 9; 456pp; English.
XX
CC The sequence given is encoded by an expression cassette which
CC consists of the hepatitis B virus (HBV) M protein (small pre-S
CC antigen, spsAg) gene precisely linked to a modified synthetic
CC vaccinia virus H6 early/late promoter. This DNA sequence was used in
CC the construction of a NYVAC recombinant expressing the HBV gene.
CC Other HBV genes were also used in the construction. These were HBV
CC L (large pre-S antigen, lpsAg) and a fusion protein composed of the
CC entire pre-S region (S12/core, S1 + S2). Each of these gene sequences
CC were inserted individually into three different sites of NYVAC
CC separated by from each other by large regions of vaccinia DNA

```

```

CC containing essential genes. NYVAC is a Copenhagen vaccine strain of
CC vaccinia virus which has been modified by deletion of six non-essential
CC regions of the genome encoding known or potential virulence factors.
CC The deletion loci were engineered as recipient loci for the insertion
CC of foreign genes. The spacing of the three inserted sequences ensured
CC that any recombination that did occur would lead to disruption of the
CC vaccinia genome and would cause unviable vaccinia virus. See also
CC AAQ35501-864.
CC (updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 281 AA;
Query Match 100.0%; Score 61; DB 13; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12
   |||||
DB 41 PLSIFSRIGDP 52

RESULT 14
AA32835
ID AA32835 standard; Protein; 281 AA.
XX
AC AA32835;
XX
DT 29-OCT-1999 (first entry)
XX
DE HBsAg PreS2-S region protein sequence.
XX
KW HBsAg; PreS2-S; recombinant antigen library; disease-related antigen;
KW multivalent antigenic polypeptide production; infection; allergen;
KW asthma; autoimmune disease; rheumatoid arthritis; diabetes; therapy;
KW multiple sclerosis; inflammatory condition; cancer; contraception;
KW immune response; hepatitis b surface antigen.
XX
OS Hepatitis b virus.
XX
PN WO9941383-A1.
XX
PD 19-AUG-1999.
XX
PF 10-FEB-1999; 99WO-US02944.
XX
PR 23-OCT-1998; 98US-0105509.
PR 11-FEB-1998; 98US-0021769.
PR 11-FEB-1998; 98US-0074294.
XX
PA (MAXY-) MAXYGEN INC.
XX
PI Bass SH, Howard R, Punnonen J, Stemmer WPC, Whalen RG;
XX
DR WPI; 1999-518452/43.
DR N-PSDB; AA210968.
XX
PT Recombinant multivalent antigenic polypeptide produced by
PT recombinant nucleic acid sequences and screening, used in vaccines
PT against e.g. infections and cancer
XX
PS Example 14; Fig 17; 153pp; English.
XX
CC This sequence is the hepatitis B virus (HBV) surface antigen (HBsAg)
CC PreS2-S region. This sequence was used to create a recombinant antigen
CC library. The library comprises recombinant nucleic acids encoding
CC antigenic polypeptides and is produced by recombination of at least two
CC forms of nucleic acid, differing by at least two nucleotides, encoding a
CC disease-related antigenic polypeptide. The library can be used to produce
CC a recombinant multivalent antigenic polypeptides of the invention, that
CC contains at least two antigenic determinants (AD) from different
CC polypeptides. The multivalent antigenic polypeptides are used in vaccines
CC to induce a protective or therapeutic response to a wide variety of
CC infectious agents (bacteria, viruses, parasites, including Plasmodium

```

CC falciparum); allergens; asthma; autoimmune disease (e.g. rheumatoid  
 CC arthritis, diabetes, multiple sclerosis); other inflammatory conditions  
 CC and cancer, also, where directed against sperm antigens, they can be used  
 CC for contraception. The multivalent peptides can be evolved to induce an  
 CC optimised immune response against a wide variety of antigens,  
 CC particularly a broad spectrum response to many different strains of a  
 CC pathogen, including strains that are likely to appear in the future.

XX  
 SQ Sequence 281 AA;

Query Match 100.0%; Score 61; DB 20; Length 281;  
 Best Local Similarity 100.0%; Pred. No. 0.0014;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
 |||||

Db 41 PLSIFSRIGDP 52

RESULT 15

ID AAR27472 standard; Protein; 389 AA.

AC AAR27472;

DT 25-MAR-2003 (updated)

DT 24-FEB-1993 (first entry)

XX  
 DE 1psAg protein.

XX Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;

KM vaccinia virus; hemorrhagic region; u promoter; NYVAC; recombinant;

KM HBV L; large pre-S antigen; 1psAg; fusion protein; pre-S region;

XX S12/core; S1; S2; Copenhagen vaccine strain; vaccinia virus;

XX virulence factor; deletion loci; recipient loci.

OS Synthetic.

XX  
 FH Key Location/Qualifiers

FT region 1..108

FT region 109..163

FT misc\_RNA /label= S2

FT 164..389

XX /label= S

XX  
 PN MO9215672-A1.

XX  
 PD 17-SEP-1992.

XX  
 PF 09-MAR-1992; 92MO-US01906.

XX  
 PR 07-MAR-1991; 91US-0666056.

XX  
 PR 11-JUN-1991; 91US-0713967.

XX  
 PR 06-MAR-1992; 92US-0847951.

XX  
 PA (VIRO-) VIROGENETICS CORP.

XX  
 PI Cox MI, De Taisne C, Francis J, Gettig RR, Johnson GP,

XX  
 PI Limbach KJ, Norton EK, Paolletti E, Perkus ME, Pincus SE;

XX  
 PI Riviere M, Tartaglia J, Taylor J;

XX  
 PS MPI; 1992-331718/40.

XX  
 DR N-PSDB; AAQ29104.

XX  
 PT Vaccine comprises recombinant, attenuated pox-virus - use for

PT  
 PT HIV, HSV, EBV, CMV, mumps etc.

XX  
 PS Disclosure; Fig 11; 456pp; English.

CC  
 CC The sequence given is encoded by an expression vector which comprises

CQ  
 CQ the hepatitis B virus (HBV) L protein (large pre-S antigen, 1psAg)

CC gene linked to the compox hemorrhagic region (u) promoter. This DNA  
 CC sequence was used in the construction of a NYVAC recombinant  
 CC expressing the HBV gene. Other HBV genes were also used in the  
 CC construction. These were HBV M protein (small pre-S antigen, spsAg)  
 CC and a fusion protein composed of the entire pre-S region (S12/core,  
 CC S1 + S2). Each of these gene sequences were inserted individually  
 CC into three different sites of NYVAC separated by from each other by  
 CC large regions of vaccinia DNA containing essential genes. NYVAC is a  
 CC Copenhagen vaccine strain of vaccinia virus which has been modified by  
 CC deletion of six non-essential regions of the genome encoding known or  
 CC potential virulence factors. The deletion loci were engineered as  
 CC recipient loci for the insertion of foreign genes. The spacing of the  
 CC three inserted sequences ensured that any recombination that did occur  
 CC would lead to disruption of the vaccinia genome and would cause  
 CC unviable vaccinia virus. See also AAQ35501-864.

XX  
 SQ Sequence 389 AA;

Query Match 100.0%; Score 61; DB 13; Length 389;  
 Best Local Similarity 100.0%; Pred. No. 0.002;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
 |||||

Db 149 PLSIFSRIGDP 160

RESULT 16

ID AAR27474 standard; Protein; 389 AA.

AC AAR27474;

DT 25-MAR-2003 (updated)

DT 24-FEB-1993 (first entry)

XX  
 DE 13L promoter/S12/core gene.

XX Hepatitis B virus; HBV; M protein; small pre-S antigen; spsAg;

KM vaccinia virus; Amsecta moorei entomopoxvirus; AMEPV; 42 KD; promoter;

KM NYVAC; recombinant; HBV L; large pre-S antigen; 1psAg; fusion protein;

XX pre-S region; S12/core; S1; S2; Copenhagen vaccine strain;

XX vaccinia virus; virulence factor; deletion loci; recipient loci.

OS Synthetic.

XX  
 FH Key Location/Qualifiers

FT region 1..108

FT region 109..163

FT Region /label= S2

FT 164..389

XX /label= S

XX  
 PN MO9215672-A1.

XX  
 PD 17-SEP-1992.

XX  
 PF 09-MAR-1992; 92MO-US01906.

XX  
 PR 07-MAR-1991; 91US-0666056.

XX  
 PR 11-JUN-1991; 91US-0713967.

XX  
 PR 06-MAR-1992; 92US-0847951.

XX  
 PA (VIRO-) VIROGENETICS CORP.

XX  
 PI Cox MI, De Taisne C, Francis J, Gettig RR, Johnson GP;

PI  
 PI Limbach KJ, Norton EK, Paolletti E, Perkus ME, Pincus SE;

PI  
 PI Riviere M, Tartaglia J, Taylor J;

XX  
 PS MPI; 1992-331718/40.

XX  
 DR N-PSDB; AAQ29106.

XX  
 PT Vaccine comprises recombinant, attenuated pox-virus - use for

PT  
 PT HIV, HSV, EBV, CMV, mumps etc.

XX  
 PS Disclosure; Fig 11; 456pp; English.

CC  
 CC The sequence given is encoded by an expression vector which comprises

CQ  
 CQ the hepatitis B virus (HBV) L protein (large pre-S antigen, 1psAg)

XX Vaccine comprises recombinant, attenuated pox-virus - use for  
PT vaccinating against viral infections such as rabies, hepatitis B,  
PT HIV, HSV, EBV, CMV, mumps etc.  
XX  
XX Disclosure; Fig 15; 456bp; English.  
XX  
XX The sequence given is encoded by an expression cassette which  
CC comprises the hepatitis B virus (HBV) L protein (large pre-S antigen,  
CC lpsag) gene which is precisely linked to the Amaceta mcore1  
CC entomopoxvirus (AmpeV) 42 kb promoter. This DNA sequence was used in  
CC the construction of a NYVAC recombinant expressing the HBV gene.  
CC Other HBV genes were also used in the construction. These were HBV M  
CC protein (small pre-S antigen, spsag) and a fusion protein composed of  
CC the entire pre-S region (S12/core, SI + S2). Each of these gene  
CC sequences were inserted individually into three different sites of  
CC NYVAC separated from each other by large regions of vaccinia DNA  
CC containing essential genes. NYVAC is a Copenhagen vaccine strain of  
CC vaccinia virus which has been modified by deletion of six non-essential  
CC regions of the genome encoding known or potential virulence factors.  
CC The deletion loci were engineered as recipient loci for the insertion  
CC of foreign genes. The spacing of the three inserted sequences ensured  
CC that any recombination that did occur would lead to disruption of the  
CC vaccinia genome and would cause unviable vaccinia virus. See also  
CC AAQ35501-864.  
CC (Updated on 25-MAR-2003 to correct FN field.)  
CC  
SQ Sequence 389 AA;  
  
Query Match 100.0%; Score 61; DB 13; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.002;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 PLSISFSRIGDP 12  
DB 149 PLSISFSRIGDP 160  
|||||  
|  
  
RESULT 17  
AAU14097  
ID AAU14097 standard; peptide; 393 AA.  
XX  
XX AAU14097;  
AC  
XX  
XX 21-NOV-2001 (first entry)  
DT  
XX  
XX Hepatitis B virus major surface antigen precursor S.  
DE  
XX  
XX Anti-retroviral; DP178-like; DP107-like; antifusogenic; antiviral;  
KW major surface antigen precursor S; HIV transmission.  
XX  
XX Hepatitis B virus subtype AYW.  
OS  
XX  
XX WO200151673-A2.  
PN  
XX  
XX 19-JUL-2001.  
PD  
XX  
XX 05-JUL-2000; 2000WO-US35727.  
PF  
XX  
XX 09-JUL-1999; 99US-0350841.  
PR  
XX  
XX (TRIM-) TRIMERIS INC.  
PA  
XX  
XX Jeffe P, Lackey JW, Erickson JB, Lawless MK, Merutka G;  
PI  
XX  
XX WPI; 2001-442157/47.  
DR  
XX  
XX Identifying a compound that inhibits the formation of or disrupts a  
PT DP107/DP178 complex, especially compounds with antifusogenic, antiviral  
PT or intracellular modulatory activity, by detecting the formation of a  
PT DP107/DP178 complex -  
XX  
XX Disclosure; Fig 35; 259bp; English.

XX  
XX The present invention relates to peptides which exhibit anti-retroviral  
CC activity. The peptides of the invention (AAU12559-AAU14009) comprise  
CC DP178-like and DP107-like peptides. The DP178 peptide corresponds  
CC to amino acids 639-673 of the transmembrane protein gp41 from human  
CC immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide  
CC corresponds to amino acids 558-595 of gp41 from HIV-1LAI. The invention  
CC also relates to a method of identifying compounds that inhibit the  
CC formation of or disrupt a DP107/DP178 complex. The method comprises  
CC detecting the formation of a DP107/DP178 complex, both in the presence  
CC or absence of a test compound, in a reaction mixture containing DP107  
CC and DP178 peptides. The method is useful for identifying compounds,  
CC including small molecule compounds, which may themselves exhibit  
CC antifusogenic, antiviral or intracellular modulatory activity. The  
CC DP178-like/DP107-like peptides are useful to inhibit human and non-human  
CC retroviral, particularly HIV, transmission to uninfected cells. The  
CC present sequence represents a peptide sequence from Hepatitis B virus  
CC subtype AYW major surface antigen precursor S.  
XX  
SQ Sequence 393 AA;  
  
Query Match 100.0%; Score 61; DB 22; Length 393;  
Best Local Similarity 100.0%; Pred. No. 0.002;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 PLSISFSRIGDP 12  
DB 153 PLSISFSRIGDP 164  
|||||  
|  
  
RESULT 18  
AAP70002  
ID AAP70002 standard; protein; 55 AA.  
XX  
XX AAP70002;  
AC  
XX  
XX 01-FEB-1991 (first entry)  
DT  
XX  
XX Sequence encoded by a later part of the Pre-S region of hepatitis  
DE B virus (HBV) beginning from the third ATG and HBV surface antigen  
DE HBSag gene including the HBV surface antigen HBSag gene.  
XX  
XX Vaccine; hepatitis; cirrhosis; cancer.  
KW  
XX  
XX Hepatitis B virus.  
OS  
XX  
XX EP241021-A.  
PN  
XX  
XX 14-OCT-1987.  
PD  
XX  
XX 08-APR-1987; 87EP-0105229.  
PF  
XX  
XX 08-APR-1986; 86JP-0079086.  
PR  
XX  
XX (GREG ) GREEN CROSS CORP.  
PA  
XX  
XX Tsujikawa M, Kobayashi K, Kawabe H, Arimura H, Nishida M, Suyama T;  
PI  
XX  
XX WPI; 1987-285906/41.  
DR  
XX  
XX N-PSDB; AAN70005.  
PD  
XX  
XX Prodn. of hepatitis B virus surface antigen - using cells  
PT transformed with recombinant DNA contg. SV40 or HBV promoter and  
PT enhancer, DNA fragment encoding Pre-S region and poly.A addn.  
PT  
XX  
XX Claim 3; Page 7; 18pp; English.  
PS  
XX  
XX The HBSag particles produced by the recombinant DNA procedure have  
CC immunological activity at least comparable to the activity of human  
CC serum-derived HBSag proteins. The particles can be used as HBV  
CC vaccine for preventing hepatitis, liver cirrhosis, liver cancer in  
CC humans. The HBSag is esp. of the sub-type adyw.  
XX

SQ Sequence 55 AA;  
 Query Match 96.7%; Score 59; DB 8; Length 55;  
 Best Local Similarity 91.7%; Pred. No. 0.00058;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12  
 |:|||||  
 41 PISSIFSRIGDP 52

Db 41 PISSIFSRIGDP 52

RESULT 19  
 AAB81133  
 ID AAB81133 standard; peptide; 64 AA.  
 XX AAB81133;  
 AC AAB81133;  
 DT 12-JUL-2001 (first entry)  
 DE Hepatitis B virus core antigen/pre-S2 construct HBC 9-87.  
 XX  
 DE Hepatitis B virus core antigen/pre-S2 construct HBC 9-87.  
 XX  
 KM Immunogen; T-helper cell response; antiviral; hepatotropic; hepatitis;  
 KM antiinflammatory; immunostimulatory; surface protein; pre-S2; HBV;  
 KM immunotherapy; core antigen.  
 XX  
 OS Hepatitis B virus.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..5  
 FT Region /note= "Hepatitis B core antigen fragment"  
 FT Region 6..61  
 FT Region /note= "Pre-S2 fragment"  
 FT Region 62..64  
 FT Region /note= "Hepatitis B core antigen fragment"  
 XX  
 PN WO200127281-A1.  
 XX  
 PD 19-APR-2001.  
 XX  
 PF 06-OCT-2000; 2000MO-GB03857.  
 XX  
 PR 08-OCT-1999; 99GB-0023902.  
 PR 30-MAR-2000; 2000GB-0007789.  
 XX  
 PA (MEDE-) MEDEVA EURO LTD.  
 PI Jones C, Bacon A, Douce G, Page M;  
 DR WPI; 2001-273776/28.  
 XX  
 PT Designing a protein immunogen useful for treating chronic viral  
 PT hepatitis comprising modifying the amino acid sequence of the immunogen  
 PT and determining whether the T-helper (Th) cells response to the  
 PT modified immunogen are Th1 or Th2 type -  
 XX  
 PS Example 5; Page 41; 94pp; English.

CC This invention relates to a method for designing protein immunogens. The  
 CC method comprises modifying the amino acid sequence of the immunogen and  
 CC determining whether the T-helper cell response to the modified immunogen  
 CC is of Th1 or Th2 type. Immunogens found to cause a T helper cell response  
 CC can be used in a pharmaceutical composition. The pharmaceutical  
 CC composition may have antiviral, hepatotropic, antiinflammatory, or  
 CC immunostimulatory activity. Determining whether an immunogen causes a Th1  
 CC or Th2 response may be useful in the treatment of diseases such as  
 CC chronic hepatitis where a Th2-dominated response fails to clear the  
 CC virus. An immunogen (isolated using the method the method of the  
 CC invention) which switches the response to Th1 may help in clearing the  
 CC virus. The method may be used to design protein immunogens which switch  
 CC the Th cell response from Th1 to Th2 and vice versa. The immunogens are  
 CC useful in the manufacture of a medicament for immunotherapeutic  
 CC application and for the treatment of diseases by immunotherapy. The

CC present sequence represents a construct consisting of a fragment of the  
 CC Hepatitis B virus (HBV) core antigen, which contains an insert of part of  
 CC the HBV surface antigen pre-S2. The peptide is used in examples of part of  
 CC illustrating the method of the invention, to show the different immune  
 CC responses caused by various modified immunogens.

SQ Sequence 64 AA;  
 Query Match 96.7%; Score 59; DB 22; Length 64;  
 Best Local Similarity 91.7%; Pred. No. 0.00068;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12  
 |:|||||  
 48 PISSIFSRIGDP 59

Db 48 PISSIFSRIGDP 59

RESULT 20  
 AAE17027  
 ID AAE17027 standard; Protein; 64 AA.  
 XX AAE17027;  
 AC AAE17027;  
 DT 18-APR-2002 (first entry)  
 DE Hepatitis B virus core antigen (HBC) full-length chimeric derivative #2.  
 XX  
 DE Hepatitis B virus core antigen (HBC) full-length chimeric derivative #2.  
 XX  
 KM Hepatitis B virus; HBV; core antigen; HBcAg; immune system; typhoid;  
 KM prophylactic; gene therapy; vaccine; hepatitis A virus; HBV; herpes;  
 KM hepatitis C virus; HCV; influenza; foot-and-mouth disease; diarrhoea;  
 KM tuberculosis; polio; rabies; acquired immunodeficiency syndrome; AIDS;  
 KM dengue fever; yellow fever; malaria; whooping cough; salmonellosis;  
 KM food poisoning; meningitis; gonorrhea; antiviral; antibacterial;  
 KM antiprotozoal; chimeric; preS2 peptide.  
 XX  
 OS Hepatitis B virus.  
 OS  
 FH Key Location/Qualifiers  
 FT Region 6..61  
 FT Region /note= "preS2 peptide"  
 XX  
 PN WO200198333-A2.  
 XX  
 PD 27-DEC-2001.  
 XX  
 PF 22-JUN-2001; 2001MO-GB02817.  
 XX  
 PR 22-JUN-2000; 2000GB-0015308.  
 PR 06-OCT-2000; 2000GB-0024544.  
 XX  
 PA (CELL-) CELUTECH PHARM LTD.  
 PI Page M, Li J, Pumps P;  
 DR WPI; 2002-098223/13.  
 XX  
 PT New proteins comprising a modified hepatitis B core antigen, useful as  
 PT a vaccine in prophylactic or therapeutic vaccination of the human or  
 PT animal body, particularly against hepatitis B virus infection -  
 XX  
 PS Disclosure; Page 25; 40pp; English.

CC The invention relates to modified proteins comprising hepatitis B virus  
 CC (HBV) core antigen (HBcAg) wherein one or more of the four arginine  
 CC repeats has been deleted and the protein comprising the C-terminal  
 CC cysteine of HBcAg. The deleted region may be replaced by an epitope  
 CC from a protein other than HBcAg, in which case the HBcAg acts as a  
 CC carrier to present the epitope to the immune system. This chimeric  
 CC protein or its nucleic acid is useful as a vaccine or in a method of  
 CC prophylactic or therapeutic vaccination of the human or animal body,  
 CC particularly against HBV. The nucleic acid encoding the protein may  
 CC be used in gene therapy or DNA vaccination protocols. The chimeric  
 CC protein or its nucleic acid may also be used as the basis of a

CC composition may have antiviral, hepatotropic, antiinflammatory, or

CC (HBV) core antigen (HBcAg) wherein one or more of the four arginine

CC repeats has been deleted and the protein comprising the C-terminal  
 CC cysteine of HBsAg. The deleted region may be replaced by an epitope  
 CC from a protein other than HBsAg, in which case the HBsAg acts as a  
 CC carrier to present the epitope to the immune system. This chimeric  
 CC protein or its nucleic acid is useful as a vaccine or in a method of  
 CC prophylactic or therapeutic vaccination of the human or animal body,  
 CC particularly against HBV. The nucleic acid encoding the protein may  
 CC be used in gene therapy or DNA vaccination protocols. The chimeric  
 CC protein or its nucleic acid may also be used as the basis of a  
 CC prophylactic vaccine against a range of diseases, e.g. HBV, hepatitis  
 CC A virus (HAV), hepatitis C virus (HCV), influenza, foot-and-mouth  
 CC disease, polio, herpes, rabies, acquired immunodeficiency syndrome  
 CC (AIDS), dengue fever, yellow fever, malaria, tuberculosis, whooping  
 CC cough, salmonellosis, typhoid, food poisoning, diarrhoea, meningitis  
 CC or gonorrhoea. The present sequence is a C-terminally truncated chimeric  
 CC derivative of Hepatitis B virus core antigen (Hbc) into which HBV  
 CC pre-S2 peptide is inserted.

XX Sequence 67 AA;

Query Match 96.7%; Score 59; DB 23; Length 67;  
 Best Local Similarity 91.7%; Pred. No. 0.00072;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSTFSRIQDP 12  
 |:|||||  
 |:|||||

DB 48 PLSSTFSRIQDP 59

RESULT 23

ID AAR15618 standard; protein; 174 AA.

XX AAR15618;

DT 25-MAR-2003 (updated)

DT 25-MAR-1992 (first entry)

DE HBsAg pre-S region subtype adyw.

XX T-cell epitope; vaccine; hepatitis B virus; antigen.

OS Synthetic.

XX WO9117768-A.

PD 26-NOV-1991.

XX 10-MAY-1991; 91WO-US03268.

XX 11-MAY-1990; 90US-0522663.

PA (SCRT-) SCRIPPS CLINIC & RE.

PI Milich DR, Thornton GB;

XX WPI; 1991-369007/50.

XX Hepatitis B virus surface antigen epitope(s) - useful as vaccines,  
 PT immunogens or diagnostic reagents

XX Claim 2; Fig 1; 91pp; English.

XX The amino acid sequence is that of a pre-S T cell epitope polypeptide  
 CC composite of the pre-S (2) regions of hepatitis B surface antigens Y  
 CC (HBsAg/Y) and d (HBsAg/d). It can be used to prime or vaccinate a host  
 CC to induce responsiveness to HBV vaccine. The T cell epitope poly-  
 CC peptides can also be used as immunogens that prime T cells that respond  
 CC to native HBsAg B cell epitope polypeptide. The T cell epitope poly-  
 CC peptides are also useful as substitutes for carrier immunogens such as  
 CC KLH and are safe, defined and T cell-active. In addition to their use  
 CC as vaccines, the polypeptides can be used as immunogens for prodn. of  
 CC antibodies. See also AAR15617-R15622.

CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 174 AA;

Query Match 96.7%; Score 59; DB 12; Length 174;  
 Best Local Similarity 91.7%; Pred. No. 0.002;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSTFSRIQDP 12  
 |:|||||  
 |:|||||

DB 160 PLSSTFSRIQDP 171

RESULT 24

ID AAR23868 standard; Protein; 174 AA.

XX AAR23868;

DT 25-MAR-2003 (updated)

DT 11-NOV-1992 (first entry)

DE Pre-S gene region translation product (2).

XX env gene; hepatitis B virus; HBV; S-protein; hydrophilic;

XX hepadnavirus; immune response; antibody; envelope.

OS Hepatitis B virus, subtype adyw.

XX EP485361-A.

PD 13-MAY-1992.

XX 25-APR-1987; 87EP-0100663.

XX 28-APR-1986; 86US-0856522.

PA (NYBL-) NY BLOOD CENT INC.

XX (CALV) CALIFORNIA INST OF TECHN.

PI Kent SBH, Neurath AR;

XX WPI; 1992-161100/20.

XX Hepatitis B pre-S peptide immunogen and vaccine - for treatment  
 PT and diagnosis of hepatitis B

XX Disclosure; Fig 2; 59pp; English.

XX The sequences given in AAR23867 - AAR23871 are amino acid sequences  
 CC deduced from sequences of the pre-S portion of the env genes  
 CC corresponding to several hepatitis B virus (HBV) subtypes. These  
 CC proteins have properties distinct from those of the S-protein in  
 CC that they have high hydrophilicity and a high percentage of charged  
 CC residues, an absence of Cys residues, the highest subtype-dependent  
 CC variability among HBV DNA gene products and little homology with  
 CC analogous sequences corresponding to nonhuman hepadnavirus. These  
 CC properties suggest that the pre-S gene coded portion of the HBV  
 CC envelope is exposed to the surface of the virion, is a target for  
 CC the host's immune response and is responsible for the host range of HBV  
 CC (limited to humans and some primates). Synthetic peptides and  
 CC antibodies against them, having predetermined specificity offer the  
 CC opportunity to explore the biological role of the pre-S protein moiety  
 CC of the HBV envelope. Portions of these amino acid sequences can be  
 CC used in a vaccine or in diagnostics for the detection of antigens and  
 CC antibodies, esp. those for the pre-S gene in sera of HBV infected  
 CC humans and certain animals, eg. chimpanzees.  
 CC (Updated on 25-MAR-2003 to correct PA field.)

XX Sequence 174 AA;

Query Match 96.7%; Score 59; DB 13; Length 174;  
 Best Local Similarity 91.7%; Pred. No. 0.002;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PLSIFSRIGDP 12  
 11:|||||  
 Db 160 PLSIFSRIGDP 171

RESULT 25  
 ABP58074  
 ID ABP58074 standard; Protein; 174 AA.  
 XX  
 AC ABP58074;  
 XX  
 DT 07-MAR-2003 (first entry)  
 XX  
 DE Hepatitis B virus subtype ayw recombinant pre-S protein.  
 XX  
 KW HBV; pre-S protein; adjuvant; vaccine; mutant; mutein.  
 XX  
 OS Hepatitis B virus.  
 OS Synthetic.  
 XX

Key Location/Qualifiers

FT Misc-difference 15 /note= "wild-type Asn substituted by His"  
 FT Misc-difference 123 /note= "wild-type Asn substituted by His"  
 FT

PN WO200294866-A1.

PD 28-NOV-2002.

PE 02-MAY-2002; 2002MO-KR00820.

PR 25-MAY-2001; 2001KR-0029002.

PA (DOBE-) DOBEEL CORP.

PI Mun H, Kim K, Yum J, Ahn B, Yi S, Seemun Y, Park J;

DR WPI; 2003-120785/11.

Modified pre-S of hepatitis B virus (HBV) enhancing immunogenicity of a general vaccine antigen, useful for the diagnosis of, prophylactic and/or therapeutic treatment of HBV infection -

Disclosure; Page 60-61; 68pp; English.

The present sequence is the protein sequence of a recombinant pre-S protein of hepatitis B virus (HBV) subtype ayw, produced by Saccharomyces cerevisiae 2805/pIL20-pre-S (adr). Cultivation of this transformant at 30 degrees C for 24-48 hours in a fed-batch system yielded recombinant pre-S, which was secreted at a level of 850 mg/l culture medium. The recombinant pre-S was purified by removing yeast cells from the broth, concentrating the cell-free broth, dialysing using an ultra-fine membrane, and separating using ion exchange and molecular size separation columns. The recombinant pre-S is modified from the native form by substitution of Asn glycosylation sites at positions 15 and 123 by histidine residues. The recombinant pre-S protein can be used as a component of an HBV vaccine for generating immunity for HBV, as an adjuvant, and in diagnostic compositions. The modifications improve the immunogenicity of the HBV S antigen, providing an improved prophylactic vaccine and a therapeutic vaccine for the treatment of chronic HBV carriers.

Sequence 174 AA;

Query Match 96.7%; Score 59; DB 24; Length 174;  
 Best Local Similarity 91.7%; Pred. No. 0.002;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PLSIFSRIGDP 12

Db 160 PLSIFSRIGDP 171  
 11:|||||

RESULT 26

AAB10596  
 ID AAB10596 standard; Protein; 347 AA.

XX  
 AC AAB10596;  
 XX

DT 08-JAN-2001 (first entry)  
 XX

DE HBV fusion protein comprising LHB and RGD.

KW Fusion protein; protein coat; virus-specific packaging signal; psi; virus protein; cell permeability; cell-specific binding site; LHB; large surface protein; core antigen; gene therapy.

OS Hepatitis b virus.

OS Synthetic.

PN WO200046376-A2.

PD 10-AUG-2000.

PE 04-FEB-2000; 2000MO-DE00363.

PR 05-FEB-1999; 99DE-1004800.

PA (HILD/) HILDT E.

PI Hildt E, Hofschneider P;

DR WPI; 2000-514959/46.

PR N-PSDB; AAA71734.

Particle for cell-specific gene delivery, useful in gene therapy, comprises nucleic acid in protein coat that includes a fusion protein of viral protein, permeability peptide and cell-binding site -

Claim 14; Fig 1; 34pp; German.

This invention describes a novel particle (A), comprising a protein coat with a fusion protein (FP), and, inside the coat, a nucleic acid (i) including the sequence for a virus-specific packaging signal (psi) and a structural gene. FP contains a virus protein (VP); a peptide (P) that mediates cell permeability and a heterologous cell-specific binding site (RGD). The invention also describes (1) producing (A) in which FP contains an LHBs (large surface protein of hepatitis B virus (HBV)) and RGD; (2) preparing (A) in which FP contains an HBV core antigen (HBcAg), (P) and RGD; (3) FP; (4) DNA encoding FP; and (5) expression vector containing the DNA of (d). The products of the invention are used in gene therapy of cells and tissues, in vivo or ex vivo. This sequence represents a fusion protein which is described in the method of the invention.

Sequence 347 AA;

Query Match 96.7%; Score 59; DB 21; Length 347;  
 Best Local Similarity 91.7%; Pred. No. 0.0043;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
 11:|||||

Db 107 PLSIFSRIGDP 118

RESULT 27

AAG66931  
 ID AAG66931 standard; Protein; 389 AA.

XX  
 AC AAG66931;  
 XX

QY 1 PLSIFSRIGDP 12

DT 19-OCT-2001 (first entry)  
 XX HBV genotype D pres1/pres2/HBsAg polypeptide.  
 DE  
 XX Hepatitis B virus; HBV; preCore; Core; pres1; pres2; HBS; HBx; HBp01;  
 KW HBsAg; antiviral; vaccine; genotype G; genotyping; HBcAg; HBsAg.  
 XX  
 OS Hepatitis B virus.  
 XX  
 PN WO200140279-A2.  
 XX  
 PD 07-JUN-2001.  
 XX  
 PF 20-NOV-2000; 2000WO-EP11526.  
 XX  
 PR 03-DEC-1999; 99EP-0870352.  
 PR 07-DEC-1999; 99US-0169287.  
 XX  
 PA (INNO-) INNOGENETICS NV.  
 XX  
 PI Stuyver L, Van Geyt C, De Gendt S;  
 DR WPI; 2001-374785/39.  
 XX  
 PT Novel isolated and/or purified hepatitis B virus polypeptide and  
 PT polynucleotide sequences that are phylogenetically different from HBV  
 PT genotype A-F molecules, useful for HBV diagnosis, prophylaxis and  
 PT therapy -  
 XX  
 PS Example 3; Fig 6; 94pp; English.  
 XX  
 CC The invention relates to the complete nucleic acid sequence of a new  
 CC human hepatitis B virus (HBV) genotype, provisionally named genotype  
 CC G. This genotype was found with a high prevalence in patients  
 CC chronically infected with HBV and residing in Europe and the USA. The  
 CC invention relates to a fully defined sequence of 3248 nucleotides as  
 CC given in specification, a sequence with 92% identity to the given  
 CC sequence, or sequence that is degenerate to the mentioned sequences.  
 CC These polynucleotides are useful for HBV genotyping. The proteins  
 CC encoded by the polynucleotides are useful for detecting antibodies in  
 CC a biological sample. Ligands that bind to the proteins and antibodies  
 CC directed against the proteins are useful for detecting the proteins  
 CC and for detecting HBcAg and HBsAg (precore precursor proteins). They  
 CC are also useful for preparing a vaccine or medicament for treating  
 CC HBV infections. The present sequence is provided in an amino acid  
 CC sequence alignment of the pres1, pres2 and HBsAg open reading frame  
 CC of the different HBV genotypes.  
 CC  
 SQ Sequence 389 AA;  
 XX  
 QY Query Match 96.7%; Score 59; DB 22; Length 389;  
 Best Local Similarity 91.7%; Pred. No. 0.0048;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 DB 149 PLSISFRRIGDP 160  
 ID AAR12394 standard; Protein; 389 AA.  
 AC AAR12394;  
 XX  
 DT 27-JAN-2003 (first entry)  
 XX  
 DE Hepatitis B virus mutant envelope protein.  
 XX  
 KW HBV; HBsAg; surface antigen; envelope; diagnosis; vaccine; mutant;  
 KW mutcin.  
 OS Hepatitis b virus.  
 XX

XX Key Location/Qualifiers  
 FH Peptide 1..108  
 FT /label= Pres1  
 FT Peptide 109..163  
 FT /label= Pres2  
 FT Protein 164..389  
 FT /label= Small\_envelope  
 XX  
 PN WO200279217-A1.  
 XX  
 PD 10-OCT-2002.  
 XX  
 PF 26-MAR-2002; 2002WO-US09227.  
 XX  
 PR 30-MAR-2001; 2001US-0821877.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Coleman PF, Mushahwar IA;  
 DR WPI; 2003-040642/03.  
 DR N-PSDB; ABV75781.  
 XX  
 PT New nucleotide sequences, useful for detecting compounds that bind to  
 PT gp41 protein or antibodies to the Hepatitis B virus surface antigen  
 PT mutant or for creating primers and probes -  
 XX  
 PS Claim 2; Fig 2; 43pp; English.  
 XX  
 CC The present sequence is the protein sequence of the entire  
 CC envelope protein of a mutant hepatitis B virus (HBV) strain  
 CC isolated from a French sample identified as 990525169. This is an  
 CC HBV subtype ayw2, genotype D strain which contains mutations leading  
 CC to 3 amino acid substitutions compared to the wild-type: Thr to Ala  
 CC 123, which affects the H166 epitope; Tyr to Leu 199, which is  
 CC outside the 'a' determinant; and Ser to Thr 207, which is also  
 CC outside the 'a' determinant. The present invention relates to the  
 CC novel HBV mutant which has a modified 'a' determinant as a result  
 CC of T123 amino acid substitution, and to methods of detecting this  
 CC mutant, and/or antibodies to the mutant, in patient samples. The  
 CC identification and detection of mutant HBV may lead to improved  
 CC vaccine development and detection systems.  
 CC  
 SQ Sequence 389 AA;  
 XX  
 QY Query Match 96.7%; Score 59; DB 24; Length 389;  
 Best Local Similarity 91.7%; Pred. No. 0.0048;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 DB 149 PLSISFRRIGDP 160  
 ID AAR12394 standard; Protein; 174 AA.  
 AC AAR12394;  
 XX  
 DT 29-AUG-1991 (first entry)  
 XX  
 DE Pre-S antigen.  
 XX  
 KW HBV; pre-SAg.  
 XX  
 OS Hepatitis B virus.  
 XX  
 PN JP03108494-A.  
 XX  
 PD 08-MAY-1991.  
 XX



PF 19-SEP-1989; 89JP-0242722.  
XX  
PR 19-SEP-1989; 89JP-0242722.  
XX  
PA (MITU ) MITSUBISHI KASEI CORP.  
XX  
XX WPI; 1991-180929/25.  
DR N-PSDB; AAQ11985.  
XX  
XX Efficient prepn. of hepatitis B virus pre-S-antigen - comprises  
PT introducing of signal peptide DNA fragment and antigen coding DNA  
PT fragment to expression vector promoter.  
XX  
XX Disclosure; Fig 1; 16pp; Japanese.  
XX  
CC The protein is expressed by a vector contg. the code for the antigen.  
CC It can be used as a vaccine for HBV. Antibodies can also be prepd.  
CC against pre-SAg and used for diagnosis of HBV.  
XX  
SQ Sequence 174 AA;  
  
Query Match 88.5%; Score 54; DB 12; Length 174;  
Best Local Similarity 83.3%; Pred. No. 0.019;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 PLSIFSRIGDP 12  
|:|||||  
Db 160 PISIFSRIGDP 171  
  
RESULT 30  
AAR15622  
ID AAR15622 standard; protein; 174 AA.  
XX  
AC AAR15622;  
XX  
DT 25-MAR-2003 (updated)  
DT 25-MAR-1992 (first entry)  
XX  
DE HBSAg pre-S region subtype adr.  
XX  
XX T-cell epitope; vaccine; hepatitis B virus; antigen.  
XX  
OS Synthetic.  
XX  
PN WO9117768-A.  
XX  
PD 28-NOV-1991.  
XX  
PF 10-MAY-1991; 91WO-US03268.  
XX  
PR 11-MAY-1990; 90US-0522663.  
XX  
PA (SCRI-) SCRIPPS CLINIC & RE.  
XX  
XX Milich DR, Thornton GB;  
PI WPI; 1991-369007/50.  
XX  
DR Hepatitis B virus surface antigen epitope(s) - useful as vaccines,  
PT immunogens or diagnostic reagents  
XX  
XX Claim 1; Fig 1; 91pp; English.  
XX  
CC The amino acid sequence is that of a pre-S T cell epitope polypeptide  
CC of the pre-S (2) region of hepatitis B surface antigen d (HBSAg/d).  
CC It can be used to prime or vaccinate a host to induce responsiveness  
CC to HBV vaccine. The T cell epitope polypeptides can also be used as  
CC immunogens that prime T cells that respond to native HBSAg B cell  
CC epitope polypeptide. The T cell epitope polypeptides are also useful  
CC as substitutes for carrier immunogens such as KdH and are safe,  
CC defined and T cell-active. In addition to their use as vaccines, the  
CC polypeptides can be used as immunogens for prodn. of antibodies. See

CC also AAR15617-R15621.  
CC (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 174 AA;  
  
Query Match 88.5%; Score 54; DB 12; Length 174;  
Best Local Similarity 83.3%; Pred. No. 0.019;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 PLSIFSRIGDP 12  
|:|||||  
Db 160 PISIFSRIGDP 171  
  
RESULT 31  
AAR23871  
ID AAR23871 standard; Protein; 174 AA.  
XX  
AC AAR23871;  
XX  
DT 25-MAR-2003 (updated)  
DT 11-NOV-1992 (first entry)  
XX  
DE Pre-S gene region translation product (5).  
XX  
XX env gene; hepatitis B virus; HBV; S-protein; hydrophilic;  
XX hepadnavirus; immune response; antibody; envelope.  
XX  
OS Hepatitis B virus, subtype adr.  
XX  
PN EP485361-A.  
XX  
PD 13-MAY-1992.  
XX  
PF 25-APR-1987; 87EP-0100663.  
XX  
PR 28-APR-1986; 86US-0856522.  
XX  
PA (NYBL-) NY BLOOD CENT INC.  
PA (CALY ) CALIFORNIA INST OF TECHN.  
XX  
PI Kent SBH, Neurath AR;  
XX  
DR WPI; 1992-161100/20.  
XX  
PT Hepatitis B pre-S peptide immunogen and vaccine - for treatment  
PT and diagnosis of hepatitis B  
XX  
PS Disclosure; Fig 2; 59pp; English.  
XX  
CC The sequences given in AAR23867 - AAR23871 are amino acid sequences  
CC deduced from sequences of the pre-S portion of the env genes  
CC corresponding to several hepatitis B virus (HBV) subtypes. These  
CC proteins have properties distinct from those of the S-protein in  
CC that they have high hydrophilicity and a high percentage of charged  
CC residues, an absence of Cys residues, the highest subtype-dependant  
CC variability among HBV DNA gene products and little homology with  
CC analogous sequences corresponding to nonhuman hepadnavirus. These  
CC properties suggest that the pre-S gene coded portion of the HBV  
CC envelope is expose to the surface of the virion, is a target for  
CC the hosts immune response and is responsible for the host range of HBV  
CC (limited to humans and some primates). Synthetic peptides and  
CC antibodies against them, having predetermined specificity offer the  
CC opportunity to explore the biological role of the pre-S protein molecy  
CC of the HBV envelope. Portions of these amino acid sequences can be  
CC used in a vaccine or in diagnostics for the detection of antigens and  
CC antibodies, esp. those for the pre-S gene in sera of HBV infected  
CC humans and certain animals, eg, chimpanzees.  
CC (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 174 AA;  
  
Query Match 88.5%; Score 54; DB 13; Length 174;

Best Local Similarity 83.3%; Pred. No. 0.019;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSISFSRIGDP 12  
|:|||||  
Db 160 PLSISFSRIGDP 171

RESULT 32  
ABP58070  
ID ABP58070 standard; Protein; 174 AA.

XX AC ABP58070;

XX DT 07-MAR-2003 (first entry)

XX DE Hepatitis B virus modified pre-S polypeptide pre-S-15m.

XX KW HBV, pre-S protein; adjuvant; vaccine; mutant; mutein.

XX OS Hepatitis B virus.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Misc-difference 15 /note="wild-type Asn substituted by His"

XX PN MO200294866-A1.

XX PD 28-NOV-2002.

XX PF 02-MAY-2002; 2002MO-KR00820.

XX PR 25-MAY-2001; 2001KR-0029002.

XX PA (DOBE-) DOBEEL CORP.

XX PI Mun H, Kim K, Yum J, Ahn B, Yi S, Seemun Y, Park J;

XX DR WPI; 2003-120785/11.

XX PT Modified pre-S of hepatitis B virus (HBV) enhancing immunogenicity of a

XX PS and/or therapeutic treatment of HBV infection -

XX CC Claim 4; Page 62-63; 68pp; English.

XX CC The present sequence is the protein sequence of a modified pre-S  
XX CC protein, pre-S-15m, of hepatitis B virus (HBV) in which the native  
XX CC asparagine residue at position 15 is substituted by a histidine  
XX CC residue. This removes one of the two glycosylation sites (the  
XX CC other is Asn-123) from the protein. The pre-S-15m protein is  
XX CC produced using a yeast (especially *Saccharomyces cerevisiae*)  
XX CC recombinant expression system. It can be used as a component of an  
XX CC HBV vaccine for generating immunity for HBV, and as an adjuvant.  
XX CC The modification improves the immunogenicity of the HBV S antigen,  
XX CC providing an improved prophylactic vaccine and a therapeutic  
XX CC vaccine for the treatment of chronic HBV carriers.

XX SQ Sequence 174 AA;

Query Match 88.5%; Score 54; DB 24; Length 174;

Best Local Similarity 83.3%; Pred. No. 0.019;

Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSISFSRIGDP 12  
|:|||||  
Db 160 PLSISFSRIGDP 171

RESULT 33

ABP58071  
ID ABP58071 standard; Protein; 174 AA.

XX AC ABP58071;

XX DT 07-MAR-2003 (first entry)

XX DE Hepatitis B virus modified pre-S polypeptide pre-S-123m.

XX KW HBV, pre-S protein; adjuvant; vaccine; mutant; mutein.

XX OS Hepatitis B virus.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Misc-difference 123 /note="wild-type Asn substituted by His"

XX PN MO200294866-A1.

XX PD 28-NOV-2002.

XX PF 02-MAY-2002; 2002MO-KR00820.

XX PR 25-MAY-2001; 2001KR-0029002.

XX PA (DOBE-) DOBEEL CORP.

XX PI Mun H, Kim K, Yum J, Ahn B, Yi S, Seemun Y, Park J;

XX DR WPI; 2003-120785/11.

XX PT Modified pre-S of hepatitis B virus (HBV) enhancing immunogenicity of a

XX PS and/or therapeutic treatment of HBV infection -

XX CC Claim 4; Page 63-64; 68pp; English.

XX CC The present sequence is the protein sequence of a modified pre-S  
XX CC protein, pre-S-123m, of hepatitis B virus (HBV) in which the native  
XX CC asparagine residue at position 123 is substituted by a histidine  
XX CC residue. This removes one of the two glycosylation sites (the  
XX CC other is Asn-15) from the protein. The pre-S-123m protein is  
XX CC produced using a yeast (especially *Saccharomyces cerevisiae*)  
XX CC recombinant expression system. It can be used as a component of an  
XX CC HBV vaccine for generating immunity for HBV, and as an adjuvant.  
XX CC The modification improves the immunogenicity of the HBV S antigen,  
XX CC providing an improved prophylactic vaccine and a therapeutic  
XX CC vaccine for the treatment of chronic HBV carriers.

XX SQ Sequence 174 AA;

Query Match 88.5%; Score 54; DB 24; Length 174;

Best Local Similarity 83.3%; Pred. No. 0.019; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSISFSRIGDP 12  
|:|||||  
Db 160 PLSISFSRIGDP 171

RESULT 34

ABP58072  
ID ABP58072 standard; Protein; 174 AA.

XX AC ABP58072;

XX DT 07-MAR-2003 (first entry)

XX DE Hepatitis B virus modified pre-S polypeptide pre-S-dm.

XX KW HBV, pre-S protein; adjuvant; vaccine; mutant; mutein.

XX OS Hepatitis B virus.

XX OS Synthetic.

```

XX Key Location/Qualifiers
FH Misc-difference 15
FT /note= "wild-type Asn substituted by His"
FT Misc-difference 123
FT /note= "wild-type Asn substituted by His"
XX
XX WO200294866-A1.
XX
XX PD 28-NOV-2002.
XX
XX PF 02-MAY-2002; 2002WO-KR00820.
XX
XX PR 25-MAY-2001; 2001KR-0029002.
XX
XX (DOBE-) DOBEEL CORP.
XX
XX PI Mun H, Kim K, Yum J, Ahn B, Yi S, Seomun Y, Park J;
XX WPI, 2003-120785/11.
XX
XX DR WPI, 2003-120785/11.
XX
XX PT Modified pre-S of hepatitis B virus (HBV) enhancing immunogenicity of a
XX general vaccine antigen, useful for the diagnosis of, prophylactic
XX and/or therapeutic treatment of HBV infection -
XX
XX PS Claim 4; Page 64-65; 68pp; English.
XX
XX CC The present sequence is the protein sequence of a modified pre-S
XX protein, pre-S-dm, of hepatitis B virus (HBV) in which the native
XX asparagine residues at positions 15 and 123 are each substituted by
XX histidine residues, thereby removing both glycosylation sites of
XX the protein. The pre-S-dm protein is produced using a yeast
XX (especially Saccharomyces cerevisiae) recombinant expression system.
XX It can be used as a component of an HBV vaccine for generating
XX immunity for HBV, and as an adjuvant. The modification improves
XX the immunogenicity of the HBV S antigen, providing an improved
XX prophylactic vaccine and a therapeutic vaccine for the treatment of
XX chronic HBV carriers.
XX
XX SQ Sequence 174 AA;
XX
XX Query Match 88.5%; Score 54; DB 24; Length 174;
XX Best Local Similarity 83.3%; Pred. No. 0.019;
XX Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 PSSSIFSRIGDP 12
XX |:|||||:|
XX 160 PISSSIFSRIGDP 171
XX
XX Db
XX
XX RESULT 35
XX ID ABP58073 standard; Protein; 174 AA.
XX
XX AC ABP58073;
XX
XX DT 07-MAR-2003 (first entry)
XX
XX DE Hepatitis B virus subtype adr recombinant pre-S protein.
XX
XX KW HBV; pre-S protein; adjuvant; vaccine; mutant; mutein.
XX
XX OS Hepatitis B virus.
XX
XX OS Synthetic.
XX
XX Key Location/Qualifiers
XX FH Misc-difference 15
XX /note= "wild-type Asn substituted by His"
XX FT Misc-difference 123
XX /note= "wild-type Asn substituted by His"
XX
XX WO200294866-A1.
XX

```

```

PD 28-NOV-2002.
XX
XX PF 02-MAY-2002; 2002WO-KR00820.
XX
XX PR 25-MAY-2001; 2001KR-0029002.
XX
XX (DOBE-) DOBEEL CORP.
XX
XX PI Mun H, Kim K, Yum J, Ahn B, Yi S, Seomun Y, Park J;
XX WPI, 2003-120785/11.
XX
XX DR WPI, 2003-120785/11.
XX
XX PT Modified pre-S of hepatitis B virus (HBV) enhancing immunogenicity of a
XX general vaccine antigen, useful for the diagnosis of, prophylactic
XX and/or therapeutic treatment of HBV infection -
XX
XX PS Disclosure; Page 59-60; 68pp; English.
XX
XX CC The present sequence is the protein sequence of a recombinant pre-S
XX protein of hepatitis B virus (HBV) subtype adr, produced by
XX Saccharomyces cerevisiae 2805/pIL20-pre-S (adr). Cultivation of
XX this transformant at 30 degrees C for 24-48 hours in a fed-batch
XX system yielded recombinant pre-S, which was secreted at a level of
XX 850 mg/l culture medium. The recombinant pre-S was purified by
XX removing yeast cells from the broth, concentrating the cell-free
XX broth, dialysing using an ultra-fine membrane, and separating using
XX ion exchange and molecular size separation columns. The
XX recombinant pre-S is modified from the native form by substitution
XX of Asn glycosylation sites at positions 15 and 123 by histidine
XX residues. The recombinant pre-S protein can be used as a component
XX of an HBV vaccine for generating immunity for HBV, as an adjuvant,
XX and in diagnostic compositions. The modifications improve the
XX immunogenicity of the HBV S antigen, providing an improved
XX prophylactic vaccine and a therapeutic vaccine for the treatment of
XX chronic HBV carriers.
XX
XX SQ Sequence 174 AA;
XX
XX Query Match 88.5%; Score 54; DB 24; Length 174;
XX Best Local Similarity 83.3%; Pred. No. 0.019;
XX Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 PSSSIFSRIGDP 12
XX |:|||||:|
XX 160 PISSSIFSRIGDP 171
XX
XX Db
XX
XX RESULT 36
XX ID AAP60163 standard; protein; 281 AA.
XX
XX AC AAP60163;
XX
XX DT 25-MAR-2003 (updated)
XX DT 02-JUL-1991 (first entry)
XX
XX DE Subtype adr hepatitis B virus surface antigen P31.
XX
XX KW Hepatitis B virus; subtype adr; surface antigen P31; vaccine;
XX
XX OS Hepatitis B virus.
XX
XX PN EP171908-A.
XX
XX PD 19-FEB-1986.
XX
XX PF 03-JUL-1985; 85RP-0304735.
XX
XX PR 03-JUN-1985; 85WO-JP00306.
XX PR 11-JUL-1984; 84WO-JP00356.
XX PR 04-SEP-1984; 84WO-JP00423.
XX PR 12-DEC-1984; 84WO-JP00585.
XX

```

PA (TAKE ) TAKEDA CHEM IND LTD.  
 PA (KIKU/) KIKUCHI M.  
 XX  
 PI Kikuchi M, Fujisawa Y, Ikeyama S, Nishimura O;  
 XX  
 DR WPI, 1986-049762/08.  
 DR N-PSDB; AAN60129.  
 XX  
 PT New non-glycosylated hepatitis-B virus surface antigen p31  
 PT protein - is prepd. by recombinant DNA methods for use in  
 PT vaccines for diagnosis, preventing and treating hepatitis-B virus  
 PT infections  
 XX  
 PS Disclosure; Fig. 1; 85bp; English.  
 XX  
 CC Subtype adr hepatitis B virus surface antigen p31 has the same biological  
 CC activity as known hepatitis B virus surface antigen small particles obtd.  
 CC from the blood of hepatitis B virus-infected individuals. The antigen is  
 CC useful in vaccines for the diagnosis, prevention and/or treatment of  
 CC hepatitis B infections in the same way as the small particles.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC  
 XX  
 SQ Sequence 281 AA;  
 Query Match 88.5%; Score 54; DB 7; Length 281;  
 Best Local Similarity 83.3%; Pred. No. 0.032;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 PLSSIFSRTGDP 12  
 |:|||||  
 Db 41 PISSIFSRTGDP 52  
 RESULT 37  
 AAP60560  
 ID AAP60560 standard; Protein; 281 AA.  
 XX  
 AC AAP60560;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 04-JUL-1991 (first entry)  
 XX  
 DE Hepatitis B virus p31 antigen subtype adr.  
 XX  
 KW Vaccine; hepatitis.  
 XX  
 OS Hepatitis B virus.  
 XX  
 PN WO8600640-A.  
 XX  
 PD 30-JAN-1986.  
 XX  
 PF 11-JUL-1984; 84WO-JP00356.  
 XX  
 PR 11-JUL-1984; 84WO-JP00356.  
 PR 03-JUN-1985; 85WO-JP00306.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 PI Kikuchi M, Fujisawa Y;  
 XX  
 DR WPI; 1986-042123/06.  
 DR N-PSDB; AAN60461.  
 XX  
 PT Hepatitis B virus surface antigen p31 from culture of  
 PT transformant - containing p31 dna at 3'-terminal promoter region.  
 PS Disclosure; Fig 1; 43bp; Japanese.  
 XX  
 CC The p31 antigen may be isolated from a transformed host in high  
 CC yields; it may be used as a vaccine against infection by the  
 CC hepatitis B virus.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC

XX  
 SQ Sequence 281 AA;  
 Query Match 88.5%; Score 54; DB 7; Length 281;  
 Best Local Similarity 83.3%; Pred. No. 0.032;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 PLSSIFSRTGDP 12  
 |:|||||  
 Db 41 PISSIFSRTGDP 52  
 RESULT 38  
 AAP60617  
 ID AAP60617 standard; Protein; 281 AA.  
 XX  
 AC AAP60617;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 25-OCT-1991 (first entry)  
 XX  
 DE C-terminal defective hepatitis B surface antigen.  
 XX  
 KW Vaccine; HBsAg.  
 XX  
 OS Hepatitis B virus.  
 XX  
 PN WO8605808-A.  
 XX  
 PD 09-OCT-1986.  
 XX  
 PF 03-APR-1985; 85WO-JP00161.  
 XX  
 PR 03-APR-1985; 85WO-JP00161.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 DR WPI; 1986-278821/42.  
 DR N-PSDB; AAN60613.  
 XX  
 PT C-terminal defective hepatitis B virus surface antigen - obtd. by  
 PT culture of transformant contg. recombinant DNA coding for  
 PT defective antigen with promoter upstream and stop codon  
 PT downstream.  
 XX  
 PS Disclosure; Fig 7; 41bp; Japanese.  
 XX  
 CC The gene product is a C-terminal defective Hepatitis B surface  
 CC antigen. The product may be incorporated into a vaccine, or Mabs  
 CC raised to it used in diagnosis of the viral infection.  
 CC The gene may be expressed in a transformed E.coli host, under the  
 CC control of a trp promoter.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 CC  
 XX  
 SQ Sequence 281 AA;  
 Query Match 88.5%; Score 54; DB 7; Length 281;  
 Best Local Similarity 83.3%; Pred. No. 0.032;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 PLSSIFSRTGDP 12  
 |:|||||  
 Db 41 PISSIFSRTGDP 52  
 RESULT 39  
 AAP70294  
 ID AAP70294 standard; protein; 281 AA.  
 XX  
 AC AAP70294;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 20-MAR-1991 (first entry)  
 DT

[illegible]

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 2003, 15:10:39 ; Search time 22 Seconds  
(without alignments)  
23.079 Million cell updates/sec

Title: US-09-830-981-2

Perfect score: 61

Sequence: 1 PLSIFSRIQDP 12

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued Patents AA: \*  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/6CTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/6backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID                | Description       |
|------------|-------|-------------|--------|----------------------|-------------------|
| 1          | 61    | 100.0       | 55     | 3 US-09-361-707-97   | Sequence 97, Appl |
| 2          | 61    | 100.0       | 164    | 4 US-09-471-573A-41  | Sequence 41, Appl |
| 3          | 61    | 100.0       | 174    | 3 US-08-480-173A-44  | Sequence 44, Appl |
| 4          | 61    | 100.0       | 174    | 3 US-08-484-408A-44  | Sequence 44, Appl |
| 5          | 61    | 100.0       | 281    | 1 US-08-105-483-214  | Sequence 214, App |
| 6          | 61    | 100.0       | 281    | 1 US-08-709-209-214  | Sequence 214, App |
| 7          | 61    | 100.0       | 281    | 1 US-08-458-101-214  | Sequence 214, App |
| 8          | 61    | 100.0       | 281    | 4 US-09-247-890-12   | Sequence 12, Appl |
| 9          | 61    | 100.0       | 281    | 4 US-09-724-969-12   | Sequence 12, Appl |
| 10         | 61    | 100.0       | 281    | 4 US-09-724-852-12   | Sequence 12, Appl |
| 11         | 61    | 100.0       | 389    | 1 US-08-105-483-216  | Sequence 216, App |
| 12         | 61    | 100.0       | 389    | 1 US-08-105-483-219  | Sequence 219, App |
| 13         | 61    | 100.0       | 389    | 1 US-08-709-209-216  | Sequence 216, App |
| 14         | 61    | 100.0       | 389    | 1 US-08-709-209-219  | Sequence 219, App |
| 15         | 61    | 100.0       | 389    | 1 US-08-458-101-216  | Sequence 216, App |
| 16         | 61    | 100.0       | 389    | 1 US-08-458-101-219  | Sequence 219, App |
| 17         | 61    | 100.0       | 389    | 3 US-08-486-099-106  | Sequence 106, App |
| 18         | 61    | 100.0       | 389    | 3 US-08-360-107A-116 | Sequence 116, App |
| 19         | 61    | 100.0       | 389    | 3 US-08-484-223B-106 | Sequence 106, App |
| 20         | 61    | 100.0       | 389    | 3 US-08-919-597-106  | Sequence 106, App |
| 21         | 61    | 100.0       | 389    | 3 US-08-475-668A-106 | Sequence 106, App |
| 22         | 61    | 100.0       | 389    | 3 US-08-485-551A-106 | Sequence 106, App |
| 23         | 61    | 100.0       | 389    | 3 US-08-471-913A-106 | Sequence 106, App |
| 24         | 61    | 100.0       | 389    | 3 US-08-485-264A-106 | Sequence 106, App |
| 25         | 61    | 100.0       | 389    | 4 US-08-474-349A-106 | Sequence 106, App |
| 26         | 61    | 100.0       | 389    | 4 US-08-470-896-106  | Sequence 106, App |
| 27         | 61    | 100.0       | 389    | 4 US-08-485-546A-106 | Sequence 106, App |

|    |    |      |     |                     |                    |
|----|----|------|-----|---------------------|--------------------|
| 28 | 59 | 96.7 | 55  | 3 US-09-361-707-93  | Sequence 93, Appl  |
| 29 | 59 | 96.7 | 55  | 3 US-09-361-707-95  | Sequence 95, Appl  |
| 30 | 59 | 96.7 | 55  | 3 US-09-361-707-96  | Sequence 96, Appl  |
| 31 | 59 | 96.7 | 174 | 2 US-08-683-262B-55 | Sequence 55, Appl  |
| 32 | 59 | 96.7 | 174 | 3 US-08-480-173A-45 | Sequence 45, Appl  |
| 33 | 59 | 96.7 | 174 | 3 US-08-484-408A-45 | Sequence 45, Appl  |
| 34 | 59 | 96.7 | 174 | 3 US-09-361-707-55  | Sequence 55, Appl  |
| 35 | 59 | 96.7 | 174 | 6 5204096-1         | Patent No. 5204096 |
| 36 | 59 | 96.7 | 174 | 6 5204096-2         | Patent No. 5204096 |
| 37 | 54 | 88.5 | 14  | 6 516485-4          | Patent No. 516485  |
| 38 | 54 | 88.5 | 15  | 3 US-09-361-707-84  | Sequence 84, Appl  |
| 39 | 54 | 88.5 | 55  | 3 US-09-361-707-85  | Sequence 85, Appl  |
| 40 | 54 | 88.5 | 55  | 3 US-09-361-707-87  | Sequence 87, Appl  |
| 41 | 54 | 88.5 | 55  | 3 US-09-361-707-89  | Sequence 89, Appl  |
| 42 | 54 | 88.5 | 55  | 3 US-09-361-707-90  | Sequence 90, Appl  |
| 43 | 54 | 88.5 | 170 | 2 US-08-683-262B-49 | Sequence 49, Appl  |
| 44 | 54 | 88.5 | 170 | 2 US-09-361-707-49  | Sequence 49, Appl  |
| 45 | 54 | 88.5 | 174 | 3 US-08-683-262B-46 | Sequence 46, Appl  |

#### ALIGNMENTS

```
RESULT 1
US-09-361-707-97
Sequence 97, Application US/09361707
Patent No. 6258937
GENERAL INFORMATION:
APPLICANT: Tong, Shuping
Li, Jisu
Mands, Jack R.
TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR
NUMBER OF SEQUENCES: 117
CORRESPONDENCE ADDRESSES:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PastsEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/361,707
FILING DATE: 27-Jul-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/683,262
FILING DATE: 18-Jul-1996
ATTORNEY/AGENT INFORMATION:
NAME: Creason, Gary L.
REGISTRATION NUMBER: 34,310
REFERENCE/DOCKET NUMBER: 00786/287003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-8906
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 55 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 97:
US-09-361-707-97
Query Match 100.0%; Score 61; DB 3; Length 55;
Best Local Similarity 100.0%; Pred. No. 0.00012;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PLSIFSRIQDP 12
Db 41 PLSIFSRIQDP 52
```

RESULT 2  
US-09-471-573A-41  
; Sequence 41, Application US/09471573A  
; Patent No. 6551820  
; GENERAL INFORMATION:  
; APPLICANT: Mason, Hugh  
; APPLICANT: Thanavajala, Yasmin  
; TITLE OF INVENTION: Expression of Immunogenic Hepatitis B Surface Antigens in Transge  
; FILE REFERENCE: 3121/1080  
; CURRENT APPLICATION NUMBER: US/09/471,573A  
; CURRENT FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: 60/113,827  
; PRIOR FILING DATE: 1998-12-23  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 41  
; LENGTH: 164  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Plant optimized pre-S (pre-S1/S2) amino acid sequence  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Plant optimized pre-S (pre-S1/S2) amino acid sequence  
US-09-471-573A-41

Query Match 100.0%; Score 61; DB 4; Length 164;  
Best Local Similarity 100.0%; Pred. No. 0.00041;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSTRIGDP 12  
DB 149 PLSISFSTRIGDP 160

RESULT 3  
US-08-480-173A-44  
; Sequence 44, Application US/08480173A  
; Patent No. 6072049  
; GENERAL INFORMATION:  
; APPLICANT: Thoma, Hans A  
; TITLE OF INVENTION: HEPATITIS B SURFACE ANTIGEN VACCINE  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Popovich & Wiles, P.A.  
; STREET: 80 S. 8th Street, Suite 1902  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/480,173A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Popovich, Thomas E  
; REGISTRATION NUMBER: 30,099  
; REFERENCE/DOCKET NUMBER: MED1003USD4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-334-8991  
; TELEFAX: 612-334-8994  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 174 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown

; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-480-173A-44

Query Match 100.0%; Score 61; DB 3; Length 174;  
Best Local Similarity 100.0%; Pred. No. 0.00044;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSTRIGDP 12  
DB 160 PLSISFSTRIGDP 171

RESULT 4  
US-08-484-408A-44  
; Sequence 44, Application US/08484408A  
; Patent No. 6117653  
; GENERAL INFORMATION:  
; APPLICANT: Thoma, Hans A  
; TITLE OF INVENTION: HEPATITIS B SURFACE ANTIGEN VACCINE  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Popovich & Wiles, P.A.  
; STREET: 80 S. 8th Street, Suite 1902  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: USA  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,408A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Popovich, Thomas E  
; REGISTRATION NUMBER: 30,099  
; REFERENCE/DOCKET NUMBER: MED1003USD4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612-334-8991  
; TELEFAX: 612-334-8994  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 174 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-484-408A-44

Query Match 100.0%; Score 61; DB 3; Length 174;  
Best Local Similarity 100.0%; Pred. No. 0.00044;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSTRIGDP 12  
DB 160 PLSISFSTRIGDP 171

RESULT 5  
US-08-105-483-214  
; Sequence 214, Application US/08105483  
; Patent No. 5494807  
; GENERAL INFORMATION:  
; APPLICANT: Paoletti, Enzo  
; TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE  
; FILE REFERENCE: STRAIN  
; NUMBER OF SEQUENCES: 462  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Curtis, Morris & Safford



ADDRESSEE: c/o William S. Frommer  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/105,483  
FILING DATE: 12-AUG-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/847,951  
FILING DATE: 06-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2400  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 214:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-105-483-214

Query Match 100.0%; Score 61; DB 1; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.00076;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFRIQDP 12  
Db 41 PLSISFRIQDP 52

RESULT 6  
US-08-709-209-214  
Sequence 214, Application US/08709209  
Patent No. 5762938  
GENERAL INFORMATION:  
APPLICANT: Paolletti, Enzo  
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE  
TITLE OF INVENTION: STRAIN  
NUMBER OF SEQUENCES: 462  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/709,209  
FILING DATE: 21-AUG-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/105,483  
FILING DATE: 12-AUG-1993  
APPLICATION NUMBER: US 07/847,951  
FILING DATE: 06-MAR-1992

ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2400  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 214:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-709-209-214

Query Match 100.0%; Score 61; DB 1; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.00076;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFRIQDP 12  
Db 41 PLSISFRIQDP 52

RESULT 7  
US-08-458-101-214  
Sequence 214, Application US/08458101  
Patent No. 5765939  
GENERAL INFORMATION:  
APPLICANT: Paolletti, Enzo  
APPLICANT: Perkus, Marion E.  
APPLICANT: Taylor, Jill  
APPLICANT: Tartaglia, James  
APPLICANT: No. 5766590cn, Elizabeth K.  
APPLICANT: Riviere, Michel  
APPLICANT: de Taisne, Charles  
APPLICANT: Limbach, Keith J.  
APPLICANT: Johnson, Gerard P.  
APPLICANT: Pincus, Steven E.  
APPLICANT: Cox, William I.  
APPLICANT: Audonnet, Jean-Christophe Francis  
APPLICANT: Gettly, Russell Robert  
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE  
TITLE OF INVENTION: STRAIN  
NUMBER OF SEQUENCES: 467  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/458,101  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2740  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 214:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 281 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-458-101-214

Query Match 100.0%; Score 61; DB 1; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.00076;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRRIGDP 12  
|||||  
Db 41 PLSISFRRIGDP 52

RESULT 8  
US-09-247-890-12  
Sequence 12, Application US/09247890  
Patent No. 6541011  
GENERAL INFORMATION:  
APPLICANT: Punnonen, Juha  
APPLICANT: Baas, Steven H.  
APPLICANT: Whalen, Robert Gerald  
APPLICANT: Howard, Russell  
APPLICANT: Stemmer, Willem P.C.  
APPLICANT: Maxygen, Inc.  
TITLE OF INVENTION: Antigen Library Immunization  
FILE REFERENCE: 018097-028710US  
CURRENT APPLICATION NUMBER: US/09/247, 890  
CURRENT FILING DATE: 1999-02-10  
EARLIER APPLICATION NUMBER: US 60/074,294  
EARLIER FILING DATE: 1998-02-11  
EARLIER APPLICATION NUMBER: US 60/105,509  
EARLIER FILING DATE: 1998-10-23  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 12  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Hepatitis B virus  
US-09-247-890-12

Query Match 100.0%; Score 61; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.00076;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRRIGDP 12  
|||||  
Db 41 PLSISFRRIGDP 52

RESULT 9  
US-09-724-969-12  
Sequence 12, Application US/09724969  
Patent No. 6569435  
GENERAL INFORMATION:  
APPLICANT: Punnonen, Juha  
APPLICANT: Baas, Steven H.  
APPLICANT: Whalen, Robert Gerald  
APPLICANT: Howard, Russell  
APPLICANT: Stemmer, Willem P.C.  
APPLICANT: Maxygen, Inc.  
TITLE OF INVENTION: Antigen Library Immunization  
FILE REFERENCE: 018097-028710US  
CURRENT APPLICATION NUMBER: US/09/724,969  
CURRENT FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: 09/247,890  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: US 60/105,509  
PRIOR FILING DATE: 1998-10-23  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 12  
LENGTH: 281  
TYPE: PRT

ORGANISM: Hepatitis B virus  
US-09-724-969-12

Query Match 100.0%; Score 61; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.00076;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRRIGDP 12  
|||||  
Db 41 PLSISFRRIGDP 52

RESULT 10  
US-09-724-852-12  
Sequence 12, Application US/09724852  
Patent No. 6576757  
GENERAL INFORMATION:  
APPLICANT: Punnonen, Juha  
APPLICANT: Baas, Steven H.  
APPLICANT: Whalen, Robert Gerald  
APPLICANT: Howard, Russell  
APPLICANT: Stemmer, Willem P.C.  
APPLICANT: Maxygen, Inc.  
TITLE OF INVENTION: Antigen Library Immunization  
FILE REFERENCE: 018097-028710US  
CURRENT APPLICATION NUMBER: US/09/724,852  
CURRENT FILING DATE: 2000-11-28  
PRIOR APPLICATION NUMBER: US/09/247,890  
PRIOR FILING DATE: 1999-02-10  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/074,294  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-02-11  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/105,509  
PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-23  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 12  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Hepatitis B virus  
US-09-724-852-12

Query Match 100.0%; Score 61; DB 4; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.00076;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRRIGDP 12  
|||||  
Db 41 PLSISFRRIGDP 52

RESULT 11  
US-08-105-483-216  
Sequence 216, Application US/08105483  
Patent No. 5494807  
GENERAL INFORMATION:  
APPLICANT: Paolletti, Enzo  
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE  
NUMBER OF SEQUENCES: 462  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Curtis, Morris & Safford  
ADDRESSEE: c/o William S. Frommer  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/105,483  
FILING DATE: 12-AUG-1993  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/847,951  
FILING DATE: 06-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2400  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 216:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-105-483-216

Query Match 100.0%; Score 61; DB 1; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIQDP 12  
Db 149 PLSISFRIQDP 160

## RESULT 12

US-08-105-483-219

Sequence 219, Application US/08105483

Patent No. 5494807

GENERAL INFORMATION:

APPLICANT: Paoletti, Enzo

TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE

TITLE OF INVENTION: STRAIN

NUMBER OF SEQUENCES: 462

CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtis, Morris & Safford

ADDRESSEE: c/o William S. Frommer

STREET: 530 Fifth Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/105,483

FILING DATE: 12-AUG-1993

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/847,951

FILING DATE: 06-MAR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Frommer, William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454310-2400

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 219:

SEQUENCE CHARACTERISTICS:

LENGTH: 389 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-105-483-219

Query Match 100.0%; Score 61; DB 1; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIQDP 12  
Db 149 PLSISFRIQDP 160

## RESULT 13

US-08-709-209-216

Sequence 216, Application US/08709209

Patent No. 5762938

GENERAL INFORMATION:

APPLICANT: Paoletti, Enzo

TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE

TITLE OF INVENTION: STRAIN

NUMBER OF SEQUENCES: 462

CORRESPONDENCE ADDRESS:

ADDRESSEE: Curtis, Morris & Safford

ADDRESSEE: c/o William S. Frommer

STREET: 530 Fifth Avenue

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/709,209

FILING DATE: 21-AUG-1996

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/105,483

FILING DATE: 12-AUG-1993

APPLICATION NUMBER: US 07/847,951

FILING DATE: 06-MAR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Frommer, William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454310-2400

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 216:

SEQUENCE CHARACTERISTICS:

LENGTH: 389 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-709-209-216

Query Match 100.0%; Score 61; DB 1; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIQDP 12  
Db 149 PLSISFRIQDP 160

## RESULT 14

US-08-709-209-219

Sequence 219, Application US/08709209

Patent No. 5762938

GENERAL INFORMATION:

APPLICANT: Paoletti, Enzo

TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE

TITLE OF INVENTION: STRAIN

US-08-709-209-219

NUMBER OF SEQUENCES: 462  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
ADDRESSEE: c/o William S. Frommer  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/709,209  
FILING DATE: 21-AUG-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/105,483  
FILING DATE: 12-AUG-1993  
APPLICATION NUMBER: US 07/847,951  
FILING DATE: 06-MAR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2400  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 219:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-709-209-219

Query Match 100.0%; Score 61; DB 1; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
|||||  
Db 149 PLSIFSRIGDP 160

RESULT 15  
US-08-458-101-216  
Sequence 216, Application US/08458101  
Patent No. 5766599  
GENERAL INFORMATION:  
APPLICANT: Paoletti, Enzo  
APPLICANT: Perkus, Marion E.  
APPLICANT: Taylor, Jill  
APPLICANT: Tartaglia, James  
APPLICANT: No. 5766599ton, Elizabeth K.  
APPLICANT: Riviere, Michel  
APPLICANT: de Taisne, Charles  
APPLICANT: Limbach, Keith J.  
APPLICANT: Johnson, Gerard P.  
APPLICANT: Pincus, Steven E.  
APPLICANT: Cox, William I.  
APPLICANT: Audonnet, Jean-Christophe  
APPLICANT: Gettig, Russell Robert  
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE  
NUMBER OF SEQUENCES: 467  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
ADDRESSEE: c/o William S. Frommer  
STREET: 530 Fifth Avenue  
CITY: New York

STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/458,101  
FILING DATE: 01-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2740  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 216:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-458-101-216

Query Match 100.0%; Score 61; DB 1; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
|||||  
Db 149 PLSIFSRIGDP 160

RESULT 16  
US-08-458-101-219  
Sequence 219, Application US/08458101  
Patent No. 5766599  
GENERAL INFORMATION:  
APPLICANT: Paoletti, Enzo  
APPLICANT: Perkus, Marion E.  
APPLICANT: Taylor, Jill  
APPLICANT: Tartaglia, James  
APPLICANT: No. 5766599ton, Elizabeth K.  
APPLICANT: Riviere, Michel  
APPLICANT: de Taisne, Charles  
APPLICANT: Limbach, Keith J.  
APPLICANT: Johnson, Gerard P.  
APPLICANT: Pincus, Steven E.  
APPLICANT: Cox, William I.  
APPLICANT: Audonnet, Jean-Christophe  
APPLICANT: Gettig, Russell Robert  
TITLE OF INVENTION: GENETICALLY ENGINEERED VACCINE  
NUMBER OF SEQUENCES: 467  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Curtis, Morris & Safford  
ADDRESSEE: c/o William S. Frommer  
STREET: 530 Fifth Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/458,101  
FILING DATE: 01-JUN-1995

CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Frommer, William S.  
REGISTRATION NUMBER: 25,506  
REFERENCE/DOCKET NUMBER: 454310-2740  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 840-3333  
TELEFAX: (212) 840-0712  
INFORMATION FOR SEQ ID NO: 219:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-458-101-219

Query Match 100.0%; Score 61; DB 1; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
DB 149 PLSIFSRIGDP 160

RESULT 17  
US-08-486-099-106  
Sequence 106, Application US/08486099  
Patent No. 6013263  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Peteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS  
TITLE OF INVENTION: B VIRUS TRANSMISSION  
NUMBER OF SEQUENCES: 209  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,099  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-031  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 106:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: protein  
MOLECULE TYPE: protein  
US-08-486-099-106

Query Match 100.0%; Score 61; DB 3; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
DB 149 PLSIFSRIGDP 160

RESULT 18  
US-08-360-107A-116  
Sequence 116, Application US/08360107A  
Patent No. 6017536  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Peteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION  
OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
TITLE OF INVENTION: TRANSMISSION  
NUMBER OF SEQUENCES: 149  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/360,107A  
FILING DATE: 20-DEC-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-013  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 116:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: protein  
MOLECULE TYPE: protein  
US-08-360-107A-116

Query Match 100.0%; Score 61; DB 3; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
DB 149 PLSIFSRIGDP 160

RESULT 19  
US-08-484-223B-106  
Sequence 106, Application US/08484223B  
Patent No. 6020459  
GENERAL INFORMATION:

APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
TITLE OF INVENTION: TRANSMISSION  
NUMBER OF SEQUENCES: 245  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,223B  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-029  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 106:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-484-223B-106

Query Match 100.0%; Score 61; DB 3; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
Db 149 PLSIFSRIGDP 160

RESULT 20  
US-08-919-597-106  
Sequence 106, Application US/08919597  
Patent No. 6054265  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION  
TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
TITLE OF INVENTION: TRANSMISSION  
NUMBER OF SEQUENCES: 273  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York

STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/919,597  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/470,896  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 106:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-919-597-106

Query Match 100.0%; Score 61; DB 3; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
Db 149 PLSIFSRIGDP 160

RESULT 21  
US-08-475-668A-106  
Sequence 106, Application US/08475668A  
Patent No. 6060065  
GENERAL INFORMATION:  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE  
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS  
TITLE OF INVENTION: TRANSMISSION  
NUMBER OF SEQUENCES: 211  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,668A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-026

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 106:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-475-668A-106

Query Match 100.0%; Score 61; DB 3; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFRIGDP 12  
Db 149 PLSISFRIGDP 160

RESULT 22  
US-08-485-551A-106  
Sequence 106, Application US/08485551A  
Patent No. 6068973  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE  
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS  
TITLE OF INVENTION: TRANSMISSION  
NUMBER OF SEQUENCES: 211  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,551A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-023  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 106:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-485-551A-106

Query Match 100.0%; Score 61; DB 3; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFRIGDP 12  
Db 149 PLSISFRIGDP 160

RESULT 23  
US-08-471-913A-106  
Sequence 106, Application US/08471913A  
Patent No. 6093794  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE  
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS  
TITLE OF INVENTION: TRANSMISSION  
NUMBER OF SEQUENCES: 214  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/471,913A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-030  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 106:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-471-913A-106

Query Match 100.0%; Score 61; DB 3; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFRIGDP 12  
Db 149 PLSISFRIGDP 160

RESULT 24  
US-08-485-264A-106  
Sequence 106, Application US/08485264A  
Patent No. 6228983  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.

APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING  
TITLE OF INVENTION: RESPIRATORY SYNCYTIAL VIRUS TRANSMISSION  
NUMBER OF SEQUENCES: 232  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,264A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-021  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 106:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-485-264A-106

Query Match 100.0%; Score 61; DB 3; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
DB 149 PLSIFSRIGDP 160

RESULT 25  
US-08-474-349A-106  
Sequence 106, Application US/08474349A  
Patent No. 6333395  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE  
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA  
NUMBER OF SEQUENCES: 517  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,349A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-024  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 106:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-474-349A-106

Query Match 100.0%; Score 61; DB 4; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
DB 149 PLSIFSRIGDP 160

RESULT 26  
US-08-470-896-106  
Sequence 106, Application US/08470896  
Patent No. 6479055  
GENERAL INFORMATION:  
APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.  
APPLICANT: Wild, Carl T.  
APPLICANT: Barney, Shawn O.  
APPLICANT: Lambert, Dennis M.  
APPLICANT: Petteway, Stephen R.  
APPLICANT: Langlois, Alphonse J.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION  
TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
NUMBER OF SEQUENCES: 273  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/470,896  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7872-020  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090



TELEFAX: (212) 869-9741/8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 106:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-470-896-106

Query Match 100.0%; Score 61; DB 4; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFRIQDP 12  
Db 149 PLSISFRIQDP 160

## RESULT 27

US-08-485-546A-106  
Sequence 106, Application US/08485546A  
Patent No. 6518013

## GENERAL INFORMATION:

APPLICANT: Bolognesi, Dani P.  
APPLICANT: Matthews, Thomas J.

APPLICANT: Wild, Carl T.

APPLICANT: Barney, Shawn O.

APPLICANT: Lambert, Dennis M.

APPLICANT: Petteley, Stephen R.

APPLICANT: Langlois, Alphonse J.

TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE

TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS

TITLE OF INVENTION: TRANSMISSION

NUMBER OF SEQUENCES: 214

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,546A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7872-028

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 106:

SEQUENCE CHARACTERISTICS:

LENGTH: 389 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-485-546A-106

Qy 1 PLSISFRIQDP 12  
Db 149 PLSISFRIQDP 160

## RESULT 28

US-09-361-707-93  
Sequence 93, Application US/09361707  
Patent No. 6258937

## GENERAL INFORMATION:

APPLICANT: Tong, Shuping

Wands, Jack R.

TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR

NUMBER OF SEQUENCES: 117

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/361,707

FILING DATE: 27-JUL-1999

PRIOR APPLICATION NUMBER: 08/683,262

FILING DATE: 18-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: Creason, Gary L.

REGISTRATION NUMBER: 34,310

REFERENCE/DOCKET NUMBER: 00786/287003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

INFORMATION FOR SEQ ID NO: 93:

SEQUENCE CHARACTERISTICS:

LENGTH: 55 amino acids

TYPE: amino acid

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 93:

Query Match 96.7%; Score 59; DB 3; Length 55;  
Best Local Similarity 91.7%; Pred. No. 0.00028;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFRIQDP 12  
Db 41 PLSISFRIQDP 52

## RESULT 29

US-09-361-707-95  
Sequence 95, Application US/09361707  
Patent No. 6258937

## GENERAL INFORMATION:

APPLICANT: Tong, Shuping

Wands, Jack R.

TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR

NUMBER OF SEQUENCES: 117

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/361,707  
FILING DATE: 27-Jul-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 06/683,262  
FILING DATE: 18-Jul-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Creason, Gary L.  
REGISTRATION NUMBER: 34,310  
REFERENCE/DOCKET NUMBER: 00786/287003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
INFORMATION FOR SEQ ID NO: 95:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 55 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 95:  
US-09-361-707-95

Query Match 96.7%; Score 59; DB 3; Length 55;  
Best Local Similarity 91.7%; Pred. No. 0.00028;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12  
Db 41 PLSIFSRIGDP 52

RESULT 30  
US-09-361-707-96  
Sequence 96, Application US/09361707  
Patent No. 6258937  
GENERAL INFORMATION:  
APPLICANT: Tong, Shuping  
L.I. Jie  
Mands, Jack R.  
TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR  
NUMBER OF SEQUENCES: 117  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/361,707  
FILING DATE: 27-Jul-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 06/683,262  
FILING DATE: 18-Jul-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Creason, Gary L.  
REGISTRATION NUMBER: 34,310  
REFERENCE/DOCKET NUMBER: 00786/287003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
INFORMATION FOR SEQ ID NO: 96:  
SEQUENCE CHARACTERISTICS:

LENGTH: 55 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 96:  
US-09-361-707-96

Query Match 96.7%; Score 59; DB 3; Length 55;  
Best Local Similarity 91.7%; Pred. No. 0.00028;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12  
Db 41 PLSIFSRIGDP 52

RESULT 31  
US-08-683-262B-55  
Sequence 55, Application US/08683262B  
Patent No. 5929220  
GENERAL INFORMATION:  
APPLICANT: Shuping Tong et al.  
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/683,262B  
FILING DATE: 18-Jul-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Frazer, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-683-262B-55

Query Match 96.7%; Score 59; DB 2; Length 174;  
Best Local Similarity 91.7%; Pred. No. 0.0011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12  
Db 160 PLSIFSRIGDP 171

RESULT 32  
US-08-480-173A-45  
Sequence 45, Application US/08480173A  
Patent No. 6072049  
GENERAL INFORMATION:  
APPLICANT: Thoma, Hans A  
TITLE OF INVENTION: HEPATITIS B SURFACE ANTIGEN VACCINE  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Popovich & Miles, P.A.  
STREET: 80 S. 8th Street, Suite 1902  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,173A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Popovich, Thomas E  
REGISTRATION NUMBER: 30,099  
REFERENCE/DOCKET NUMBER: MED1003USD4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-334-8991  
TELEFAX: 612-334-8994  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-480-173A-45

Query Match 96.7%; Score 59; DB 3; Length 174;  
Best Local Similarity 91.7%; Pred. No. 0.0011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PSSISFRRIGDP 12  
Db 160 PSSISFRRIGDP 171

RESULT 33  
US-08-484-408A-45  
Sequence 45, Application US/08484408A  
Patent No. 6117653  
GENERAL INFORMATION:  
APPLICANT: Thoma, Hans A  
TITLE OF INVENTION: HEPATITIS B SURFACE ANTIGEN VACCINE  
NUMBER OF SEQUENCES: 56  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Popovich & Miles, P.A.  
STREET: 80 S. 8th Street, Suite 1902  
CITY: Minneapolis  
STATE: MN  
COUNTRY: USA  
ZIP: 55402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,408A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Popovich, Thomas E  
REGISTRATION NUMBER: 30,099  
REFERENCE/DOCKET NUMBER: MED1003USD4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 612-334-8991  
TELEFAX: 612-334-8994  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:

LENGTH: 174 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-484-408A-45

Query Match 96.7%; Score 59; DB 3; Length 174;  
Best Local Similarity 91.7%; Pred. No. 0.0011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PSSISFRRIGDP 12  
Db 160 PSSISFRRIGDP 171

RESULT 34  
US-09-361-707-55  
Sequence 55, Application US/09361707  
Patent No. 6258937  
GENERAL INFORMATION:  
APPLICANT: Tong, Shuping  
L1, Jisu  
Wande, Jack R.  
TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR  
NUMBER OF SEQUENCES: 117  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/361,707  
FILING DATE: 27-Jul-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Creason, Gary L.  
REGISTRATION NUMBER: 34,310  
REFERENCE/DOCKET NUMBER: 00786/287003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 55:

US-09-361-707-55

Query Match 96.7%; Score 59; DB 3; Length 174;  
Best Local Similarity 91.7%; Pred. No. 0.0011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PSSISFRRIGDP 12  
Db 160 PSSISFRRIGDP 171

RESULT 35  
US-09-520-4096-1  
Patent No. 5204096  
APPLICANT: NEURATH, ALEXANDER R.; KENT, B.H.

TITLE OF INVENTION: PRE-S GENE CODED PEPTIDE HEPATITIS B  
IMMUNOGENS, VACCINES, DIAGNOSTICS, AND SYNTHETIC LIPID VESICLE  
CARRIERS  
NUMBER OF SEQUENCES: 36  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/338,028  
FILING DATE: 14-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 698,499  
FILING DATE: 05-FEB-1985  
APPLICATION NUMBER: 587,090  
FILING DATE: 07-MAR-1984  
SEQ ID NO: 1:  
LENGTH: 174  
5204096-1

Query Match 96.7%; Score 59; DB 6; Length 174;  
Best Local Similarity 91.7%; Pred. No. 0.0011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRTGDP 12  
|:|||||  
Db 160 PLSISFRTGDP 171

RESULT 36  
5204096-2  
PATENT NO. 5204096  
APPLICANT: NEURATH, ALEXANDER R.; KENT, B.H.  
TITLE OF INVENTION: PRE-S GENE CODED PEPTIDE HEPATITIS B  
IMMUNOGENS, VACCINES, DIAGNOSTICS, AND SYNTHETIC LIPID VESICLE  
CARRIERS  
NUMBER OF SEQUENCES: 36  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/338,028  
FILING DATE: 14-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 698,499  
FILING DATE: 05-FEB-1985  
APPLICATION NUMBER: 587,090  
FILING DATE: 07-MAR-1984  
SEQ ID NO: 2:  
LENGTH: 174  
5204096-2

Query Match 96.7%; Score 59; DB 6; Length 174;  
Best Local Similarity 91.7%; Pred. No. 0.0011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRTGDP 12  
|:|||||  
Db 160 PLSISFRTGDP 171

RESULT 37  
5164485-4  
PATENT NO. 5164485  
APPLICANT: FUJISAWA, YUKIO; ITOH, YASUAKI; NISHIMURA, OSAMU  
FUJII, TOMOYO  
TITLE OF INVENTION: MODIFIED HEPATITIS B VIRUS SURFACE  
ANTIGEN P31 AND PRODUCTION THEREOF  
NUMBER OF SEQUENCES: 22  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/547,948  
FILING DATE: 03-JUL-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 898,425  
FILING DATE: 20-AUG-1986  
SEQ ID NO: 4:  
LENGTH: 14  
5164485-4

Query Match 88.5%; Score 54; DB 6; Length 14;

Best Local Similarity 83.3%; Pred. No. 0.00052;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSISFRTGDP 12  
|:|||||  
Db 1 PLSISFRTGDP 12

RESULT 38  
US-09-361-707-84  
SEQUENCE 84, Application US/09361707  
PATENT NO. 6258937  
GENERAL INFORMATION:  
APPLICANT: Tong, Shuping  
Li, Jisu  
Mands, Jack R.  
TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR  
NUMBER OF SEQUENCES: 117  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for windows version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/361,707  
FILING DATE: 27-Jul-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Creason, Gary L.  
REGISTRATION NUMBER: 34,310  
REFERENCE/DOCKET NUMBER: 00786/287003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
INFORMATION FOR SEQ ID NO: 84:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 55 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 84:  
US-09-361-707-84

Query Match 88.5%; Score 54; DB 3; Length 55;  
Best Local Similarity 83.3%; Pred. No. 0.0025;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSISFRTGDP 12  
|:|||||  
Db 41 PLSISFRTGDP 52

RESULT 39  
US-09-361-707-85  
SEQUENCE 85, Application US/09361707  
PATENT NO. 6258937  
GENERAL INFORMATION:  
APPLICANT: Tong, Shuping  
Li, Jisu  
Mands, Jack R.  
TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR  
NUMBER OF SEQUENCES: 117  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street

CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/361,707  
FILING DATE: 27-Jul-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/683,262  
FILING DATE: 18-Jul-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Creason, Gary L.  
REGISTRATION NUMBER: 34,310  
REFERENCE/DOCKET NUMBER: 00786/287003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
INFORMATION FOR SEQ ID NO: 85:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 55 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 85:  
US-09-361-707-85

Query Match 88.5%; Score 54; DB 3; Length 55;  
Best Local Similarity 83.3%; Pred. No. 0.0025;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSISFSRIGDP 12  
|:|||||  
DB 41 PLSISFSRTGDP 52

RESULT 40  
US-09-361-707-87  
Sequence 87, Application US/09361707  
Patent No. 6258937  
GENERAL INFORMATION:  
APPLICANT: Tong, Shuping  
L1, Jisu  
Wands, Jack R.  
TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR  
NUMBER OF SEQUENCES: 117  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/361,707  
FILING DATE: 27-Jul-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/683,262  
FILING DATE: 18-Jul-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Creason, Gary L.  
REGISTRATION NUMBER: 34,310  
REFERENCE/DOCKET NUMBER: 00786/287003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906  
INFORMATION FOR SEQ ID NO: 87:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 55 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 87:  
US-09-361-707-87

Query Match 88.5%; Score 54; DB 3; Length 55;  
Best Local Similarity 83.3%; Pred. No. 0.0025;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSISFSRIGDP 12  
|:|||||  
DB 41 PLSISFSRTGDP 52

Search completed: November 6, 2003, 15:13:49  
Job time : 22 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 2003, 15:12:49 ; Search time 29 Seconds

(without alignments)  
71.069 Million cell updates/sec

Title: US-09-830-981-2

Perfect score: 61

Sequence: 1 PLSISFRIGDP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues 644079

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID                | Description       |
|------------|-------|-------------|--------|----|-------------------|-------------------|
| 1          | 61    | 100.0       | 12     | 15 | US-10-226-956-284 | Sequence 284, App |
| 2          | 61    | 100.0       | 12     | 15 | US-10-077-555-9   | Sequence 9, Appl  |
| 3          | 61    | 100.0       | 12     | 15 | US-10-211-088-302 | Sequence 302, App |
| 4          | 61    | 100.0       | 12     | 15 | US-09-247-890-12  | Sequence 12, Appl |
| 5          | 59    | 96.7        | 174    | 9  | US-09-818-066-55  | Sequence 55, Appl |
| 6          | 59    | 96.7        | 389    | 10 | US-09-821-877-2   | Sequence 2, Appl  |
| 7          | 54    | 88.5        | 170    | 9  | US-09-818-066-49  | Sequence 49, Appl |
| 8          | 54    | 88.5        | 174    | 9  | US-09-818-066-46  | Sequence 46, Appl |
| 9          | 54    | 88.5        | 174    | 9  | US-09-818-066-48  | Sequence 48, Appl |
| 10         | 54    | 88.5        | 174    | 9  | US-09-818-066-56  | Sequence 56, Appl |
| 11         | 54    | 88.5        | 281    | 10 | US-09-247-890-10  | Sequence 10, Appl |
| 12         | 54    | 88.5        | 400    | 15 | US-10-209-264-3   | Sequence 3, Appl  |
| 13         | 52    | 85.2        | 174    | 9  | US-09-818-066-50  | Sequence 50, Appl |
| 14         | 52    | 85.2        | 174    | 9  | US-09-818-066-58  | Sequence 58, Appl |
| 15         | 52    | 85.2        | 389    | 16 | US-10-169-668-6   | Sequence 6, Appl  |

|    |      |      |     |    |                     |                   |
|----|------|------|-----|----|---------------------|-------------------|
| 16 | 49   | 80.3 | 174 | 9  | US-09-818-066-53    | Sequence 53, Appl |
| 17 | 46   | 75.4 | 174 | 9  | US-09-818-066-45    | Sequence 45, Appl |
| 18 | 46   | 75.4 | 174 | 9  | US-09-818-066-47    | Sequence 47, Appl |
| 19 | 44   | 72.1 | 174 | 9  | US-09-818-066-37    | Sequence 37, Appl |
| 20 | 44   | 72.1 | 174 | 9  | US-09-818-066-54    | Sequence 54, Appl |
| 21 | 43   | 70.5 | 174 | 9  | US-09-818-066-57    | Sequence 57, Appl |
| 22 | 40   | 65.6 | 64  | 9  | US-09-917-340-85    | Sequence 85, Appl |
| 23 | 40   | 65.6 | 64  | 15 | US-10-157-031-118   | Sequence 118, App |
| 24 | 40   | 65.6 | 174 | 9  | US-09-818-066-42    | Sequence 42, Appl |
| 25 | 40   | 65.6 | 174 | 9  | US-09-818-066-62    | Sequence 62, Appl |
| 26 | 38   | 62.3 | 174 | 9  | US-09-818-066-38    | Sequence 38, Appl |
| 27 | 38   | 62.3 | 174 | 9  | US-09-818-066-40    | Sequence 40, Appl |
| 28 | 37   | 60.7 | 174 | 9  | US-09-818-066-41    | Sequence 41, Appl |
| 29 | 37   | 60.7 | 174 | 9  | US-09-818-066-59    | Sequence 59, Appl |
| 30 | 37   | 60.7 | 217 | 12 | US-10-291-190-42    | Sequence 42, Appl |
| 31 | 36.5 | 59.8 | 335 | 12 | US-09-995-938A-8    | Sequence 8, Appl  |
| 32 | 36.5 | 59.8 | 335 | 12 | US-09-995-938A-10   | Sequence 10, Appl |
| 33 | 36   | 59.0 | 55  | 9  | US-09-879-257A-45   | Sequence 45, Appl |
| 34 | 36   | 59.0 | 174 | 9  | US-09-818-066-35    | Sequence 35, Appl |
| 35 | 36   | 59.0 | 174 | 9  | US-09-818-066-36    | Sequence 36, Appl |
| 36 | 36   | 59.0 | 174 | 9  | US-09-818-066-52    | Sequence 52, Appl |
| 37 | 36   | 59.0 | 174 | 16 | US-10-169-668-4     | Sequence 4, Appl  |
| 38 | 35   | 57.4 | 96  | 9  | US-09-864-761-39796 | Sequence 39796, A |
| 39 | 35   | 57.4 | 110 | 15 | US-10-106-698-6697  | Sequence 6697, Ap |
| 40 | 35   | 57.4 | 173 | 12 | US-10-205-979-48    | Sequence 48, Appl |
| 41 | 35   | 57.4 | 173 | 14 | US-10-051-643-207   | Sequence 207, App |
| 42 | 35   | 57.4 | 174 | 9  | US-09-818-066-39    | Sequence 39, Appl |
| 43 | 35   | 57.4 | 400 | 12 | US-10-224-999A-3461 | Sequence 3461, Ap |
| 44 | 35   | 57.4 | 613 | 15 | US-10-156-761-13218 | Sequence 13218, A |
| 45 | 35   | 57.4 | 690 | 9  | US-09-815-242-12460 | Sequence 12460, A |

## ALIGNMENTS

RESULT 1  
US-10-226-956-284  
; Sequence 284, Application US/10226956  
; Publication No. US20030060399A1  
; GENERAL INFORMATION:  
; APPLICANT: Brophy, Colleen  
; APPLICANT: Komalavilaa, Padmini  
; APPLICANT: Pantich, Alyssa  
; APPLICANT: Joshi, Lokesh  
; APPLICANT: Seal, Brandon L.  
; TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES  
; FILE REFERENCE: ASU-1061-US  
; CURRENT APPLICATION NUMBER: US/10/226,956  
; CURRENT FILING DATE: 2002-08-23  
; PRIOR APPLICATION NUMBER: 60/314,535  
; PRIOR FILING DATE: 2001-08-23  
; NUMBER OF SEQ ID NOS: 320  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 284  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic peptide  
US-10-226-956-284

Query Match 100.0%; Score 61; DB 15; Length 12;  
Best Local Similarity 100.0%; Pred. No. 7.1e-05;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIGDP 12  
|||||  
Db 1 PLSISFRIGDP 12

RESULT 2  
US-10-077-555-9  
; Sequence 9, Application US/10077555

```
/ Publication No. US20030077289A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Rong-fu
/ TITLE OF INVENTION: Use of Cell-Penetrating Peptides to Generate Antitumor Immunity
/ FILE REFERENCE: P02373US1/10200806
/ CURRENT APPLICATION NUMBER: US/10/077,555
/ PRIOR FILING DATE: 2002-02-15
/ PRIOR APPLICATION NUMBER: US 60/268,687
/ PRIOR FILING DATE: 2001-02-15
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 9
/ LENGTH: 12
/ TYPE: PRT
/ ORGANISM: Hepatitis B virus
US-10-077-555-9

Query Match          100.0%; Score 61; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 7,1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PLSISFRRIGDP 12
        |||||
        1 PLSISFRRIGDP 12

Db
1 PLSISFRRIGDP 12

RESULT 3
US-10-211-088-302
/ Sequence 302, Application US/10211088
/ Publication No. US20030104479A1
/ GENERAL INFORMATION:
/ APPLICANT: Bright, Gary R.
/ APPLICANT: Premkumar, D. David
/ APPLICANT: Chen, Yih-Tai
/ TITLE OF INVENTION: No. US20030104479A1 Fusion Proteins And Assays For Molecular Bi
/ FILE REFERENCE: 01-1022-US
/ CURRENT APPLICATION NUMBER: US/10/211,088
/ PRIOR FILING DATE: 2002-10-15
/ PRIOR APPLICATION NUMBER: 60/309,395
/ PRIOR FILING DATE: 2001-08-01
/ PRIOR APPLICATION NUMBER: 60/341,589
/ PRIOR FILING DATE: 2001-12-13
/ NUMBER OF SEQ ID NOS: 366
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 302
/ LENGTH: 12
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Protein-derived transport peptide
US-10-211-088-302

Query Match          100.0%; Score 61; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 7,1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PLSISFRRIGDP 12
        |||||
        1 PLSISFRRIGDP 12

Db
1 PLSISFRRIGDP 12

RESULT 4
US-09-247-890-12
/ Sequence 12, Application US/09247890
/ Publication No. US20020198162A1
/ GENERAL INFORMATION:
/ APPLICANT: Punnonen, Juha
/ APPLICANT: Baas, Steven H.
/ APPLICANT: Whalen, Robert Gerald
/ APPLICANT: Howard, Russell
/ APPLICANT: Stemmer, Willem P.C.
/ APPLICANT: Maxygen, Inc.
/ TITLE OF INVENTION: Antigen Library Immunization
```

```
/ FILE REFERENCE: 018097-028710US
/ CURRENT APPLICATION NUMBER: US/09/247,890
/ CURRENT FILING DATE: 1999-02-10
/ EARLIER APPLICATION NUMBER: US 60/074,294
/ EARLIER FILING DATE: 1998-02-11
/ EARLIER APPLICATION NUMBER: US 60/105,509
/ EARLIER FILING DATE: 1998-10-23
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: Patentin Ver. 2.0
/ SEQ ID NO 12
/ LENGTH: 281
/ TYPE: PRT
/ ORGANISM: Hepatitis B virus
US-09-247-890-12

Query Match          100.0%; Score 61; DB 10; Length 281;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PLSISFRRIGDP 12
        |||||
        41 PLSISFRRIGDP 52

Db
41 PLSISFRRIGDP 52

RESULT 5
US-09-818-066-55
/ Sequence 55, Application US/09818066
/ Patent No. US20020032307A1
/ GENERAL INFORMATION:
/ APPLICANT: Shuping Tong et al.
/ TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
/ NUMBER OF SEQUENCES: 75
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Fish & Richardson P.C.
/ STREET: 225 Franklin Street
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: U.S.A.
/ ZIP: 02110-2804
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/818,066
/ FILING DATE: 27-Mar-2001
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/683,262
/ FILING DATE: 18-Jul-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Fraser, Janis K.
/ REGISTRATION NUMBER: 31,819
/ REFERENCE/DOCKET NUMBER: 00786/287002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 542-5070
/ TELEFAX: (617) 542-8906
/ TELEX: 200154
/ INFORMATION FOR SEQ ID NO: 55:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 174 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 55:
US-09-818-066-55

Query Match          96.7%; Score 59; DB 9; Length 174;
Best Local Similarity 91.7%; Pred. No. 0.0031;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 PLSISFRRIGDP 12
```



Db 160 PISSTFSRIGDP 171

|||||

RESULT 6

US-09-821-877-2

Sequence 2, Application US/09821877

Patent No. US20020177124A1

GENERAL INFORMATION:

APPLICANT: Abbott Laboratories

APPLICANT: Coleman, Paul F.

APPLICANT: Mushahwar, Isa K.

TITLE OF INVENTION: Hepatitis B Virus Surface Antigen Mutant

TITLE OF INVENTION: And Methods Of Detection Thereof

FILE REFERENCE: 6794 US 01

CURRENT APPLICATION NUMBER: US/09/821,877

CURRENT FILING DATE: 2000-03-03

NUMBER OF SEQ ID NOS: 8

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 389

TYPE: PRT

ORGANISM: Hepatitis B Virus

US-09-821-877-2

Query Match

Best Local Similarity 96.7%; Score 59; DB 10; Length 389;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSTFSRIGDP 12

Db 149 PISSTFSRIGDP 160

|||||

RESULT 7

US-09-818-066-49

Sequence 49, Application US/09818066

Patent No. US20020032307A1

GENERAL INFORMATION:

APPLICANT: Shuping Tong et al.

TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR

NUMBER OF SEQUENCES: 75

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/818,066

FILING DATE: 27-Mar-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/683,262

FILING DATE: 18-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: Fraser, Janis K.

REGISTRATION NUMBER: 31,819

REFERENCE/DOCKET NUMBER: 00786/287002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 49:

SEQUENCE CHARACTERISTICS:

LENGTH: 170 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 49:

US-09-818-066-49

Query Match

Best Local Similarity 88.5%; Score 54; DB 9; Length 170;

Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSSTFSRIGDP 12

Db 156 PISSTFSRIGDP 167

|||||

RESULT 8

US-09-818-066-46

Sequence 46, Application US/09818066

Patent No. US20020032307A1

GENERAL INFORMATION:

APPLICANT: Shuping Tong et al.

TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR

NUMBER OF SEQUENCES: 75

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/818,066

FILING DATE: 27-Mar-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/683,262

FILING DATE: 18-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: Fraser, Janis K.

REGISTRATION NUMBER: 31,819

REFERENCE/DOCKET NUMBER: 00786/287002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:

LENGTH: 174 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 46:

US-09-818-066-46

Query Match

Best Local Similarity 88.5%; Score 54; DB 9; Length 174;

Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSSTFSRIGDP 12

Db 160 PISSTFSRIGDP 171

|||||

RESULT 9

US-09-818-066-48

Sequence 48, Application US/09818066

Patent No. US20020032307A1

GENERAL INFORMATION:

APPLICANT: Shuping Tong et al.

TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 48:  
US-09-818-066-48  
Query Match 88.5%; Score 54; DB 9; Length 174;  
Best Local Similarity 83.3%; Pred. No. 0.026; 1; Indels 0; Gaps 0;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 PLSIFSRIGDP 12  
Db 160 PLSIFSRIGDP 171  
RESULT 10  
US-09-818-066-56  
Sequence 56; Application US/09818066  
Patent No. US20020032307A1  
GENERAL INFORMATION:  
APPLICANT: Shuping Tong et al.  
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262

FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 56:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 56:  
US-09-818-066-56  
Query Match 88.5%; Score 54; DB 9; Length 174;  
Best Local Similarity 100.0%; Pred. No. 0.026; 0; Indels 0; Gaps 0;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 PLSIFSRIGD 11  
Db 160 PLSIFSRIGD 170  
RESULT 11  
US-09-247-890-10  
Sequence 10; Application US/09247890  
Publication No. US20020198162A1  
GENERAL INFORMATION:  
APPLICANT: Punnonen, Juha  
APPLICANT: Baes, Steven H.  
APPLICANT: Whalen, Robert Gerald  
APPLICANT: Howard, Russell  
APPLICANT: Stemmer, William P.C.  
APPLICANT: Maxygen, Inc.  
TITLE OF INVENTION: Antigen Library Immunization  
FILE REFERENCE: 018097-028710US  
CURRENT APPLICATION NUMBER: US/09/247,890  
EARLIER FILING DATE: 1999-02-10  
EARLIER APPLICATION NUMBER: US 60/074,294  
EARLIER FILING DATE: 1998-02-11  
EARLIER APPLICATION NUMBER: US 60/105,509  
EARLIER FILING DATE: 1998-10-23  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 10  
LENGTH: 281  
TYPE: PRT  
ORGANISM: Hepatitis B virus  
US-09-247-890-10  
Query Match 88.5%; Score 54; DB 10; Length 281;  
Best Local Similarity 83.3%; Pred. No. 0.044; 1; Indels 0; Gaps 0;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 PLSIFSRIGDP 12  
Db 41 PLSIFSRIGDP 52  
RESULT 12  
US-10-209-264-3  
Sequence 3; Application US/10209264  
Publication No. US20030003111A1  
GENERAL INFORMATION:  
APPLICANT: Oon, Chong Jin  
Lim, Gek Keow  
Zhao, Yi  
Chen, Wei Ning  
TITLE OF INVENTION: A MUTANT HUMAN HEPATITIS B VIRAL STRAIN AND

USSES THEREOF

NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ladac & Parry  
STREET: 26 West 61 Street  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10023

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/209,264  
FILING DATE: 31-Jul-2002  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/SG98/00046  
FILING DATE: 19-Jan-1998

ATTORNEY/AGENT INFORMATION:  
NAME: Mass, Clifford J.  
REGISTRATION NUMBER: 30,086  
REFERENCE/DOCKET NUMBER: U-013109-7  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 708-1800

INFORMATION FOR SEQ ID NO: 3:  
LENGTH: 400 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-209-264-3

Query Match 88.5%; Score 54; DB 15; Length 400;  
Best Local Similarity 83.3%; Pred. No. 0.065;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
Db 160 PLSIFSRIGDP 171

RESULT 13  
US-09-818-066-50  
Sequence 50, Application US/09818066  
Patent No. US20020032307A1

GENERAL INFORMATION:  
APPLICANT: Shuping Tong et al.  
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154

INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 50:  
US-09-818-066-50

Query Match 85.2%; Score 52; DB 9; Length 174;  
Best Local Similarity 90.9%; Pred. No. 0.062;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 ISSIFSRIGDP 12  
Db 161 ISSIFSRIGDP 171

RESULT 14  
US-09-818-066-58  
Sequence 58, Application US/09818066  
Patent No. US20020032307A1

GENERAL INFORMATION:  
APPLICANT: Shuping Tong et al.  
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154

INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 58:  
US-09-818-066-58

QY 2 LSSIFSRIGDP 12  
 DB 161 ISSIFSRIGDP 171

RESULT 15  
 US-10-169-668-6  
 ; Sequence 6, Application US/10169668  
 ; Publication No. US20030129202A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BIOMERIEUX  
 ; APPLICANT: INSERM  
 ; TITLE OF INVENTION: MUTATED HEPATITIS B VIRUS, ITS NUCLEIC AND PROTEIN  
 ; FILE REFERENCE: IFB 99 INS HBVM  
 ; CURRENT APPLICATION NUMBER: US/10/169,668  
 ; CURRENT FILING DATE: 2002-07-08  
 ; NUMBER OF SEQ ID NOS: 6  
 ; SOFTWARE: Patentln Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 389  
 ; TYPE: PRT  
 ; ORGANISM: mutated hepatitis B virus mHBV  
 US-10-169-668-6

Query Match 85.2%; Score 52; DB 16; Length 389;  
 Best Local Similarity 90.9%; Pred. No. 0.15;  
 Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSSIFSRIGDP 12  
 DB 150 ISSIFSRIGDP 160

RESULT 16  
 US-09-818-066-53  
 ; Sequence 53, Application US/09818066  
 ; Patent No. US20020032307A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shuping Tong et al.  
 ; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
 ; NUMBER OF SEQUENCES: 75  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02110-2804  
 ; TELEFAX: (617) 542-8906  
 ; TELEPHONE: (617) 542-5070  
 ; TELECOMMUNICATION INFORMATION:  
 ; REFERENCE/DOCKET NUMBER: 00786/287002  
 ; FILING DATE: 27-Mar-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/683,262  
 ; FILING DATE: 18-JUL-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Frazer, Janis K.  
 ; REGISTRATION NUMBER: 31,819  
 ; REFERENCE/DOCKET NUMBER: 00786/287002  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 542-5070  
 ; TELEFAX: (617) 542-8906  
 ; TELEX: 200154  
 ; INFORMATION FOR SEQ ID NO: 53:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 174 amino acids  
 ; TYPE: amino acid

TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 53:  
 US-09-818-066-53

Query Match 80.3%; Score 49; DB 9; Length 174;  
 Best Local Similarity 81.8%; Pred. No. 0.23;  
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSSIFSRIGDP 12  
 DB 161 ISSIFSRIGDP 171

RESULT 17  
 US-09-818-066-45  
 ; Sequence 45, Application US/09818066  
 ; Patent No. US20020032307A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shuping Tong et al.  
 ; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
 ; NUMBER OF SEQUENCES: 75  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02110-2804  
 ; TELEFAX: (617) 542-8906  
 ; TELEPHONE: (617) 542-5070  
 ; TELECOMMUNICATION INFORMATION:  
 ; REFERENCE/DOCKET NUMBER: 00786/287002  
 ; FILING DATE: 27-Mar-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/683,262  
 ; FILING DATE: 18-JUL-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Frazer, Janis K.  
 ; REGISTRATION NUMBER: 31,819  
 ; REFERENCE/DOCKET NUMBER: 00786/287002  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 542-5070  
 ; TELEFAX: (617) 542-8906  
 ; TELEX: 200154  
 ; INFORMATION FOR SEQ ID NO: 45:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 174 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 45:  
 US-09-818-066-45

Query Match 75.4%; Score 46; DB 9; Length 174;  
 Best Local Similarity 75.0%; Pred. No. 0.83;  
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LSSIFSRIGDP 12  
 DB 160 PISISSRIGDP 171

RESULT 18  
 US-09-818-066-47  
 ; Sequence 47, Application US/09818066  
 ; Patent No. US20020032307A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shuping Tong et al.

TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 47:  
US-09-818-066-47

Query Match 75.4%; Score 46; DB 9; Length 174;  
Best Local Similarity 75.0%; Pred. No. 0.83;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12  
Db 160 PLSISSRTGDP 171

RESULT 19  
US-09-818-066-37  
Sequence 37, Application US/09818066  
Patent No. US20020032307A1  
GENERAL INFORMATION:  
APPLICANT: Shuping Tong et al.  
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262

FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 37:  
US-09-818-066-37

Query Match 72.1%; Score 44; DB 9; Length 174;  
Best Local Similarity 81.8%; Pred. No. 2;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LSSIFSRIGDP 12  
Db 161 ISSISSRTGDP 171

RESULT 20  
US-09-818-066-54  
Sequence 54, Application US/09818066  
Patent No. US20020032307A1  
GENERAL INFORMATION:  
APPLICANT: Shuping Tong et al.  
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 54:  
US-09-818-066-54

Query Match 72.1%; Score 44; DB 9; Length 174;  
Best Local Similarity 75.0%; Pred. No. 2;

Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 PLSISFRIQDP 12  
|:|||||  
Db 160 PLSISFRTIDP 171

RESULT 21  
US-09-818-066-57  
; Sequence 57, Application US/09818066  
; Patent No. US20020032307A1  
; GENERAL INFORMATION:  
; APPLICANT: Shuping Tong et al.  
; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
; NUMBER OF SEQUENCES: 75  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/818,066  
; FILING DATE: 27-Mar-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/683,262  
; FILING DATE: 18-JUL-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Frazer, Janis K.  
; REGISTRATION NUMBER: 31,819  
; REFERENCE/DOCKET NUMBER: 00786/287002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154

## INFORMATION FOR SEQ ID NO: 57:

SEQUENCE CHARACTERISTICS:  
; LENGTH: 174 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 57:  
US-09-818-066-57

Query Match 70.5%; Score 43; DB 9; Length 174;  
Best Local Similarity 72.7%; Pred. No. 3;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSISFRIQDP 11  
|:|||||  
Db 160 PLSISFRTIDP 170

RESULT 22  
US-09-917-340-85  
; Sequence 85, Application US/09917340  
; Patent No. US20020090369A1  
; GENERAL INFORMATION:  
; APPLICANT: Murphy, Christopher J.  
; APPLICANT: McAnulty, Jonathan F.  
; APPLICANT: Reid, Ted W.  
; TITLE OF INVENTION: Transplant Media  
; FILE REFERENCE: TPLANT-06468  
; CURRENT APPLICATION NUMBER: US/09/917,340  
; CURRENT FILING DATE: 2001-07-29  
; PRIOR APPLICATION NUMBER: 60/221,632

; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: 60/249,602  
; PRIOR FILING DATE: 2000-11-17  
; PRIOR APPLICATION NUMBER: 60/290,932  
; PRIOR FILING DATE: 2001-05-15  
; NUMBER OF SEQ ID NOS: 96  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 85  
; LENGTH: 64  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-917-340-85

Query Match 65.6%; Score 40; DB 9; Length 64;  
Best Local Similarity 58.3%; Pred. No. 3.7;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 PLSISFRIQDP 12  
|:|||||  
Db 17 PLPGVFGIGIDP 28

RESULT 23  
US-10-157-031-118  
; Sequence 118, Application US/10157031  
; Publication No. US20030108890A1  
; GENERAL INFORMATION:  
; APPLICANT: Baranova, A. V.  
; APPLICANT: Yankovsky, N. K.  
; APPLICANT: Kozlov, A. P.  
; APPLICANT: Lobashev, A. V.  
; APPLICANT: Krukovskaya, L. L.  
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences  
; FILE REFERENCE: 2760-103  
; CURRENT APPLICATION NUMBER: US/10/157,031  
; CURRENT FILING DATE: 2002-05-30  
; NUMBER OF SEQ ID NOS: 415  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 118  
; LENGTH: 64  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-157-031-118

Query Match 65.6%; Score 40; DB 15; Length 64;  
Best Local Similarity 58.3%; Pred. No. 3.7;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 PLSISFRIQDP 12  
|:|||||  
Db 17 PLPGVFGIGIDP 28

RESULT 24  
US-09-818-066-42  
; Sequence 42, Application US/09818066  
; Patent No. US20020032307A1  
; GENERAL INFORMATION:  
; APPLICANT: Shuping Tong et al.  
; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
; NUMBER OF SEQUENCES: 75  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 42:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 42:  
US-09-818-066-42

Query Match 65.6%; Score 40; DB 9; Length 174;  
Best Local Similarity 66.7%; Pred. No. 11;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 USSIFSRIGDP 12  
DB 160 PISSTESRTGDP 171

RESULT 25  
US-09-818-066-62  
Sequence 62, Application US/09818066  
Patent No. US2002003307A1  
GENERAL INFORMATION:  
APPLICANT: Shuping Tong et al.  
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 62:  
US-09-818-066-62

Query Match 65.6%; Score 40; DB 9; Length 174;  
Best Local Similarity 63.6%; Pred. No. 11;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 USSIFSRIGDP 12  
DB 161 ISSVSTGDP 171

RESULT 26  
US-09-818-066-38  
Sequence 38, Application US/09818066  
Patent No. US2002003307A1  
GENERAL INFORMATION:  
APPLICANT: Shuping Tong et al.  
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 38:  
US-09-818-066-38

Query Match 62.3%; Score 38; DB 9; Length 174;  
Best Local Similarity 63.6%; Pred. No. 26;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 USSIFSRIGDP 12  
DB 161 ISSVSTGDP 171

RESULT 27  
US-09-818-066-40  
Sequence 40, Application US/09818066  
Patent No. US2002003307A1  
GENERAL INFORMATION:  
APPLICANT: Shuping Tong et al.  
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR

NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 40:  
US-09-818-066-40  
Query Match 62.3%; Score 38; DB 9; Length 174;  
Best Local Similarity 63.6%; Pred. No. 26;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 LSIFSRIGDP 12  
DB 161 ISSILSKTGP 171  
RESULT 28  
US-09-818-066-41  
Sequence 41, Application US/09818066  
Patent No. US20020032307A1  
GENERAL INFORMATION:  
APPLICANT: Shuping Tong et al.  
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 41:  
US-09-818-066-41  
Query Match 62.3%; Score 38; DB 9; Length 174;  
Best Local Similarity 63.6%; Pred. No. 26;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 LSIFSRIGDP 12  
DB 161 ISSILSKTGP 171  
RESULT 29  
US-09-818-066-59  
Sequence 59, Application US/09818066  
Patent No. US20020032307A1  
GENERAL INFORMATION:  
APPLICANT: Shuping Tong et al.  
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 59:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 59:  
US-09-818-066-59  
Query Match 60.7%; Score 37; DB 9; Length 174;  
Best Local Similarity 72.7%; Pred. No. 40;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;



QY 2 ISSIFSRIGDP 12  
 :|||||  
 Db 161 ISSIFSRITDP 171

# RESULT 30

US-10-291-190-42  
 ; Sequence 42, Application US/10291190  
 ; Publication No. US20030171549A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Structural Genomics, Inc.  
 ; APPLICANT: Park, Frances  
 ; APPLICANT: Gajiwala, Ketan S.  
 ; APPLICANT: Buchanan, Sean Grant  
 ; APPLICANT: Sauder, J. Michael  
 ; TITLE OF INVENTION: CRYSTALS AND STRUCTURES OF YIIM PROTEINS  
 ; FILE REFERENCE: 52498-20009.00  
 ; CURRENT APPLICATION NUMBER: US/10/291,190  
 ; CURRENT FILING DATE: 2002-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/337,769  
 ; PRIOR FILING DATE: 2001-11-09  
 ; NUMBER OF SEQ ID NOS: 58  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 42  
 ; LENGTH: 217  
 ; TYPE: PRT  
 ; ORGANISM: A. tumefaciens (15887448)  
 US-10-291-190-42

Query Match 60.7%; Score 37; DB 12; Length 217;  
 Best Local Similarity 50.0%; Pred. No. 51;  
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 PLUSIFSRIGDP 12  
 :|||:|||||  
 Db 129 PCATLSARIGDP 140

# RESULT 31

US-09-995-938A-8  
 ; Sequence 8, Application US/09995938A  
 ; Publication No. US20030150026A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: JOANNE CHORY  
 ; APPLICANT: ZHIYONG WANG  
 ; TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID  
 ; TITLE OF INVENTION: HORMONE ACTION IN PLANTS  
 ; FILE REFERENCE: SALKINS.046A  
 ; CURRENT APPLICATION NUMBER: US/09/995,938A  
 ; CURRENT FILING DATE: 2001-11-27  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 8  
 ; LENGTH: 335  
 ; TYPE: PRT  
 ; ORGANISM: ARABIDOPSIS THALIANA  
 US-09-995-938A-8

Query Match 59.8%; Score 36.5; DB 12; Length 335;  
 Best Local Similarity 60.0%; Pred. No. 1e+02;  
 Matches 9; Conservative 1; Mismatches 2; Indels 3; Gaps 1;

QY 1 PLUSIF--SRIGDP 12  
 :|||:|||||  
 Db 130 PSSSFPSPSRVGD 144

# RESULT 32

US-09-995-938A-10  
 ; Sequence 10, Application US/09995938A  
 ; Publication No. US20030150026A1  
 ; GENERAL INFORMATION:

; APPLICANT: JOANNE CHORY  
 ; APPLICANT: ZHIYONG WANG  
 ; TITLE OF INVENTION: GENES INVOLVED IN BRASSINOSTEROID  
 ; TITLE OF INVENTION: HORMONE ACTION IN PLANTS  
 ; FILE REFERENCE: SALKINS.046A  
 ; CURRENT APPLICATION NUMBER: US/09/995,938A  
 ; CURRENT FILING DATE: 2001-11-27  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 10  
 ; LENGTH: 335  
 ; TYPE: PRT  
 ; ORGANISM: ARABIDOPSIS THALIANA  
 US-09-995-938A-10

Query Match 59.8%; Score 36.5; DB 12; Length 335;  
 Best Local Similarity 60.0%; Pred. No. 1e+02;  
 Matches 9; Conservative 1; Mismatches 2; Indels 3; Gaps 1;

QY 1 PLUSIF--SRIGDP 12  
 :|||:|||||  
 Db 130 PSSSFPSPSRVGD 144

# RESULT 33

US-09-879-257A-45  
 ; Sequence 45, Application US/09879257A  
 ; Patent No. US20020081690A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: YAMAMOTO, SACHIKO  
 ; APPLICANT: HANADA, TOSHIRO  
 ; APPLICANT: SHIRO, MINORO  
 ; APPLICANT: KOBATAKE, SHINZO  
 ; TITLE OF INVENTION: HYBRID ENZYMES AND USE THEREOF  
 ; FILE REFERENCE: 55986 (70281)  
 ; CURRENT APPLICATION NUMBER: US/09/879,257A  
 ; CURRENT FILING DATE: 2001-06-12  
 ; NUMBER OF SEQ ID NOS: 56  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 45  
 ; LENGTH: 55  
 ; TYPE: PRT  
 ; ORGANISM: Hepatitis B Virus  
 US-09-879-257A-45

Query Match 59.0%; Score 36; DB 9; Length 55;  
 Best Local Similarity 63.6%; Pred. No. 18;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ISSIFSRIGDP 12  
 :||||:|||||  
 Db 42 ISSISARTGDP 52

# RESULT 34

US-09-818-066-35  
 ; Sequence 35, Application US/09818066  
 ; Patent No. US20020032307A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Shuping Tong et al.  
 ; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
 ; NUMBER OF SEQUENCES: 75  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: Massachusetts  
 ; COUNTRY: U.S.A.  
 ; ZIP: 02110-2804  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 35:  
US-09-818-066-35  
Query Match 59.0%; Score 36; DB 9; Length 174;  
Best Local Similarity 63.6%; Pred. No. 61;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 LSSIFSRIGDP 12  
Db 161 ISSISARTGDP 171  
RESULT 35  
US-09-818-066-36  
Sequence 36, Application US/09818066  
Patent No. US20020032307A1  
GENERAL INFORMATION:  
APPLICANT: Shuping Tong et al.  
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 36:  
US-09-818-066-36  
Query Match 59.0%; Score 36; DB 9; Length 174;  
Best Local Similarity 63.6%; Pred. No. 61;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 LSSIFSRIGDP 12  
Db 161 ISSISARTGDP 171  
RESULT 36  
US-09-818-066-52  
Sequence 52, Application US/09818066  
Patent No. US20020032307A1  
GENERAL INFORMATION:  
APPLICANT: Shuping Tong et al.  
TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR  
NUMBER OF SEQUENCES: 75  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/818,066  
FILING DATE: 27-Mar-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/683,262  
FILING DATE: 18-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Fraser, Janis K.  
REGISTRATION NUMBER: 31,819  
REFERENCE/DOCKET NUMBER: 00786/287002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 174 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 52:  
US-09-818-066-52  
Query Match 59.0%; Score 36; DB 9; Length 174;  
Best Local Similarity 63.6%; Pred. No. 61;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 LSSIFSRIGDP 12  
Db 161 ISSISARTGDP 171  
RESULT 37  
US-10-169-668-4  
Sequence 4, Application US/10169668  
Publication No. US20030129202A1  
GENERAL INFORMATION:  
APPLICANT: BIOMERIEUX

```

; APPLICANT: INSEEM
; TITLE OF INVENTION: MUTATED HEPATITIS B VIRUS, ITS NUCLEIC AND PROTEIN
; FILE REFERENCE: IEB 99 INS HBVM
; CURRENT APPLICATION NUMBER: US/10/169,668
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 174
; TYPE: PRT
; ORGANISM: mutated hepatitis B virus MHV
US-10-169-668-4

Query Match          59.0%; Score 36; DB 16; Length 174;
Best Local Similarity 63.6%; Pred. No. 61;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LSIFSRIGDP 12
Db 161 ISSISARTGDP 171

RESULT 38
US-09-864-761-39796
; Sequence 39796, Application US/09664761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
```

```

; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39796
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005691.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 16
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 17
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 21
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 18
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 19
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 29
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 21
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 24
; OTHER INFORMATION: EST_HUMAN HIT: AM028249.1, EVALUATE 1.00e-04
US-09-864-761-39796

Query Match          57.4%; Score 35; DB 9; Length 96;
Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LSIFSRIGDP 12
Db 64 LVSDPFRVGDGP 74

RESULT 39
US-10-106-698-6697
; Sequence 6697, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patentin Ver. 3.0
; SEQ ID NO 6697
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (13)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (27)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (37)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (57)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (59)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (85)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (86)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
```

```

; LOCATION: (97)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (103)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6697

```

```

Query Match          57.4%; Score 35; DB 15; Length 110;
Best Local Similarity 50.0%; Pred. No. 57;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

```

QY      1 PLSIFSRIGDP 12
        | : | : | : |
Db      23 PSCITXPARTGDP 34

```

```

RESULT 40
US-10-205-979-48
; Sequence 48; Application US/10205979
; Publication No. US20030147861A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; APPLICANT: Abernethy, Nevin
; TITLE OF INVENTION: Compounds and Methods for the Modulation
; FILE REFERENCE: 11000.1063U
; CURRENT APPLICATION NUMBER: US/10/205,979
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/308,446
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Mycobacterium vaccae
US-10-205-979-48

```

```

Query Match          57.4%; Score 35; DB 12; Length 173;
Best Local Similarity 54.5%; Pred. No. 94;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 PLSIFSRIGDP 11
        | | | : | : |
Db      127 PLSMFQVGD 137

```

Search completed: November 6, 2003, 15:18:06  
 Job time : 29 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 6, 2003, 15:09:39 / Search time 21 Seconds  
(without alignments)  
54.954 Million cell updates/sec

Title: US-09-830-981-2  
Perfect score: 61  
Sequence: 1 PLSIFSRIGDP 12

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID    | Description        |
|------------|-------|-------------|--------|----------|--------------------|
| 1          | 61    | 100.0       | 378    | 2 S41869 | surface antigen -  |
| 2          | 61    | 100.0       | 389    | 1 SAVLAH | large surface anti |
| 3          | 61    | 100.0       | 389    | 2 S20749 | surface antigen -  |
| 4          | 59    | 96.7        | 389    | 1 SAVLBH | large surface anti |
| 5          | 59    | 96.7        | 389    | 1 SAVLAJ | large surface anti |
| 6          | 59    | 96.7        | 389    | 1 SAVLAI | large surface anti |
| 7          | 59    | 96.7        | 389    | 2 S47407 | surface antigen -  |
| 8          | 59    | 96.7        | 389    | 2 S32202 | large surface anti |
| 9          | 54    | 88.5        | 382    | 2 T13474 | large surface anti |
| 10         | 54    | 88.5        | 384    | 2 T13469 | large surface anti |
| 11         | 54    | 88.5        | 389    | 2 S20745 | surface antigen -  |
| 12         | 54    | 88.5        | 400    | 1 SAVLA  | large surface anti |
| 13         | 54    | 88.5        | 400    | 2 S35528 | surface antigen -  |
| 14         | 54    | 88.5        | 445    | 2 S43492 | surface antigen -  |
| 15         | 47    | 77.0        | 389    | 2 S67506 | large surface anti |
| 16         | 45    | 73.8        | 378    | 2 S41870 | surface antigen -  |
| 17         | 45    | 73.8        | 389    | 2 S20753 | surface antigen -  |
| 18         | 43    | 70.5        | 358    | 2 B81194 | conserved hypotet  |
| 19         | 43    | 70.5        | 358    | 2 A81831 | hypothetical prote |
| 20         | 41    | 67.2        | 502    | 2 C87709 | ubiquitome biosynt |
| 21         | 40    | 65.6        | 389    | 1 SAVLCP | large surface anti |
| 22         | 40    | 65.6        | 389    | 2 S41871 | surface antigen -  |
| 23         | 39    | 63.9        | 399    | 2 AF2349 | hypothetical prote |
| 24         | 39    | 63.9        | 557    | 4 S42326 | hypothetical large |
| 25         | 38    | 62.3        | 278    | 2 AF2895 | hypothetical prote |
| 26         | 38    | 62.3        | 315    | 2 A86807 | hypothetical prote |
| 27         | 38    | 62.3        | 389    | 1 SAVLJ1 | large surface anti |
| 28         | 38    | 62.3        | 389    | 1 SAVLJ2 | large surface anti |
| 29         | 38    | 62.3        | 449    | 2 F86852 | amino acid permeas |

|    |      |      |      |          |                    |
|----|------|------|------|----------|--------------------|
| 30 | 38   | 62.3 | 537  | 2 S67434 | probable sterol O- |
| 31 | 38   | 62.3 | 603  | 2 H69121 | hypothetical prote |
| 32 | 38   | 62.3 | 1786 | 1 H71527 | probable exinucle  |
| 33 | 37   | 60.7 | 134  | 2 A82455 | hypothetical prote |
| 34 | 37   | 60.7 | 134  | 2 G82477 | hypothetical prote |
| 35 | 37   | 60.7 | 229  | 2 A97370 | hypothetical prote |
| 36 | 37   | 60.7 | 229  | 2 A12587 | conserved hypotet  |
| 37 | 37   | 60.7 | 333  | 2 A90956 | probable integrase |
| 38 | 37   | 60.7 | 333  | 2 F85804 | integrase for prop |
| 39 | 37   | 60.7 | 422  | 2 T50516 | retinoid x recepto |
| 40 | 37   | 60.7 | 495  | 2 T21115 | hypothetical prote |
| 41 | 37   | 60.7 | 533  | 1 S74401 | hypothetical prote |
| 42 | 37   | 60.7 | 582  | 2 D84362 | hypothetical prote |
| 43 | 37   | 60.7 | 1037 | 2 E84723 | hypothetical prote |
| 44 | 36.5 | 59.8 | 335  | 2 G86326 | protein F18014.7 l |
| 45 | 36   | 59.0 | 230  | 2 A75260 | hypothetical prote |

## ALIGNMENTS

RESULT 1  
S41869  
surface antigen - hepatitis B virus (subtype ayw, patient A) (fragment)  
N/Alternate names: envelope protein, Hbs antigen  
N/Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (ml  
C/Species: hepatitis B virus, HBV  
A/Variety: subtype ayw, patient A  
C/Date: 06-Jan-1995 #sequence\_revision 01-Nov-1996 #text\_change 26-Aug-1999  
R/Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.  
submitted to the EMBL Data Library, January 1994  
A/Description: Sequence analysis of HBV genomes isolated from patients with HBeAg chroni  
A/Reference number: S41869  
A/Accession: S41869  
A/Molecule type: DNA  
A/Residues: 1-378 <LA1>  
A/Cross-references: EMBL:X77309; NID:9452611; PIDN:CA54515.1; PID:9452612  
A/Experimental source: subtype ayw, patient A  
C/Genetics:  
A/Gene: S  
C/Superfamily: hepatitis B virus surface antigen  
C/Keywords: surface antigen  
F/1-378/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>  
F/1-108/Domain: pre-S1 domain #status predicted <PRE1>  
F/109-378/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <F  
F/109-163/Domain: pre-S2 domain #status predicted <PRE2>  
F/164-378/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 100.0% Score 61; DB 2; Length 378;  
Best Local Similarity 100.0%; Pred. No. 0.0012;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
|||||  
DB 149 PLSIFSRIGDP 160

RESULT 2  
SAVLAH  
large surface antigen - hepatitis B virus (subtype ayw)  
N/Contains: major surface antigen; middle surface antigen  
C/Species: hepatitis B virus, HBV  
C/Date: 18-Dec-1981 #sequence\_revision 18-Dec-1981 #text\_change 07-May-1999  
C/Accession: A03703; J02064; P00591  
R/Galibert, F.; Mandart, B.; Fitoussi, F.; Tiollais, P.; Charney, P.  
Nucleur 281, 646-650, 1979  
A/Title: Nucleotide sequence of the hepatitis B virus genome (subtype ayw) in E. coli.  
A/Reference number: A93214; MUID:81012091; PMID:399327  
A/Accession: A03703  
A/Molecule type: DNA  
A/Residues: 1-389 <GAL>  
A/Cross-references: GB:J02203

R:Norder, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouge, A.M.; Mushahwar, I.K.; Magnus, J. Gen. Virol. 74, 1341-1348, 1993  
 A>Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin  
 A:Reference number: JQ2044; PMID:93329382; PMID:8336122  
 A:Accession: JQ2064  
 A:Molecule type: DNA  
 A:Residues: 164-389 <NOR>  
 A:Experimental source: genogroup D, subtype ayw3  
 R:Norder, H.; Courouge, A.M.; Magnus, L.O.  
 J. Gen. Virol. 73, 3141-3145, 1992  
 A>Title: Molecular basis of hepatitis B virus serotype variations within the four major  
 A:Reference number: PQ0453; PMID:93107848; PMID:1469353  
 A:Accession: PQ0591  
 A:Molecule type: DNA  
 A:Residues: 264-343 <NO2>  
 A:Experimental source: subtype ayw3, KI1  
 C:Genetics:  
 A:Gene: pre-S1/pre-S2/S  
 C:Superfamily: hepatitis B virus surface antigen  
 C:Keywords: glycoprotein; surface antigen  
 F:109-389/Product: middle surface antigen (gene S) #status predicted <MSA>  
 F:164-389/Product: major surface antigen (gene S) #status predicted  
 F:4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 61; DB 1; Length 389;  
 Best Local Similarity 100.0%; Pred. No. 0.0012;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12  
 DB 149 PLSSIFSRIGDP 160

## RESULT 3

Surface antigen - hepatitis B virus (subtype ayw, patient CI)  
 N:Alternate names: envelope protein; HBs antigen  
 N:Contains: major surface antigen; middle surface antigen  
 C:Species: hepatitis B virus, HBV  
 A:Variety: subtype ayw, patient CI  
 C:Date: 20-Feb-1995 #sequence\_revision 01-Nov-1996 #text\_change 26-Aug-1999  
 C:Accession: S20749  
 R:Pal, M.B.; Mazzoleni, A.P.; Balestrieri, A.; Mellis, A.; Porru, A.  
 Submitted to the EMBL Data Library, March 1992  
 A:Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negative  
 A:Reference number: S20745  
 A:Accession: S20749  
 A:Molecule type: DNA  
 A:Residues: 1-389 <LA1>  
 A:Cross-references: EMBL:X65258; NID:G59434; PID:G59435  
 A:Experimental source: subtype ayw, patient CI  
 A>Note: typical methionine for the beginning of pre-S2 domain at position 109 is missing  
 C:Genetics:  
 A:Gene: S  
 C:Superfamily: hepatitis B virus surface antigen  
 C:Keywords: surface antigen  
 F:1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>  
 F:164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 100.0%; Score 61; DB 2; Length 389;  
 Best Local Similarity 100.0%; Pred. No. 0.0012;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12  
 DB 149 PLSSIFSRIGDP 160

RESULT 4  
 SAYLBH  
 Large surface antigen - hepatitis B virus (subtype ayw, strain PH8320)  
 N:Contains: major surface antigen; middle surface antigen  
 C:Species: hepatitis B virus, HBV

A>Note: host Homo sapiens (man)  
 C:Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 28-Jul-2000  
 C:Accession: A03704; PQ0585  
 R:Bicho, V.; Pushko, P.; Drellina, D.; Pumpen, P.; Gren, E.  
 FEBS Lett. 185, 208-212, 1985  
 A>Title: Subtype ayw variant of hepatitis B virus: DNA primary structure analysis.  
 A:Reference number: A05237; PMID:85204397; PMID:3996597  
 A:Accession: A03704  
 A:Molecule type: DNA  
 A:Residues: 1-389 <BIC>  
 A:Cross-references: GB:X02496; NID:962280; PIDN:CAB41701.1; PID:g4704321  
 R:Norder, H.; Courouge, A.M.; Magnus, L.O.  
 J. Gen. Virol. 73, 3141-3145, 1992  
 A>Title: Molecular basis of hepatitis B virus serotype variations within the four major  
 A:Reference number: PQ0453; PMID:93107848; PMID:1469353  
 A:Accession: PQ0585  
 A:Molecule type: DNA  
 A:Residues: 264-343 <NOR>  
 A:Experimental source: subtype ayw2, Tav  
 C:Genetics:  
 A:Gene: pre-S1/pre-S2/S  
 C:Superfamily: hepatitis B virus surface antigen  
 C:Keywords: glycoprotein; surface antigen  
 F:109-389/Product: middle surface antigen (gene S) #status predicted <MSA>  
 F:164-389/Product: major surface antigen (gene S) #status predicted  
 F:4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 96.7%; Score 59; DB 1; Length 389;  
 Best Local Similarity 91.7%; Pred. No. 0.0029;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12  
 DB 149 PLSSIFSRIGDP 160

## RESULT 5

Large surface antigen - hepatitis B virus (subtype adyw)  
 N:Contains: major surface antigen; middle surface antigen  
 C:Species: hepatitis B virus, HBV  
 C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 15-Nov-1996  
 C:Accession: A93217; A03703  
 R:Pasak, M.; Goto, T.; Gilbert, W.; Zink, B.; Schaller, H.; Mackay, P.; Leadbetter, G.;  
 Nature 287, 575-579, 1979  
 A>Title: Hepatitis B virus genes and their expression in E. coli.  
 A:Reference number: A93217; PMID:8102115; PMID:399329  
 A:Accession: A93217  
 A:Molecule type: DNA  
 A:Residues: 1-389 <PAS>  
 A:Cross-references: GB:J02202  
 C:Genetics:  
 A:Gene: pre-S1/pre-S2/S  
 C:Superfamily: hepatitis B virus surface antigen  
 C:Keywords: glycoprotein; surface antigen  
 F:109-389/Product: middle surface antigen (gene S) #status predicted <MSA>  
 F:164-389/Product: major surface antigen (gene S) #status predicted  
 F:4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 96.7%; Score 59; DB 1; Length 389;  
 Best Local Similarity 91.7%; Pred. No. 0.0029;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12  
 DB 149 PLSSIFSRIGDP 160

RESULT 6  
 SAYLBH  
 Large surface antigen - hepatitis B virus (strain alpha)  
 N:Contains: major surface antigen; middle surface antigen  
 C:Species: hepatitis B virus, HBV

C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 16-Jul-1999  
C/Accession: B34773; JQ2082  
R/Tong, S.; Li, J.; Valtvicki, L.; Trepo, C.  
Virology 176, 596-603, 1990  
A/Title: Active hepatitis B virus replication in the presence of anti-HBe is associated  
A/Reference number: A34773; MUID:90266476; PMID:2345966  
A/Accession: B34773  
A/Status: translation not shown  
A/Molecule type: DNA  
A/Residues: 1-389 <TON>  
A/Cross-references: EMBL:M32138; NID:g329667; PIDN:AAA45502.1; PID:g329669  
R/Norder, H.; Hammes, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,  
J. Gen. Virol. 74, 1341-1348, 1993  
A/Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin  
A/Reference number: JQ2044; MUID:93329382; PMID:8336122  
A/Accession: JQ2082  
A/Molecule type: DNA  
A/Residues: 164-389 <NOR>  
A/Experimental source: genogroup D, subtype ayw2, strain HBV-alpha  
C/Genetics:  
A/Introns: 111/3  
A/Genes: pre-S1/pre-S2/S  
C/Superfamily: hepatitis B virus surface antigen  
C/Keywords: glycoprotein; surface antigen (gene pre-S2/S) #status predicted <DSa>  
F/109-389/Product: middle surface antigen (gene S) #status predicted <MSA>  
F/164-389/Product: major surface antigen (gene S) #status predicted  
F/4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 96.7%; Score 59; DB 1; Length 389;  
Best Local Similarity 91.7%; Pred. No. 0.0029;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRRIGDP 12  
DB 149 PLSISFRRIGDP 160

RESULT 7  
S47407  
surface antigen - hepatitis B virus (subtype ayw4)  
N/Alternate names: envelope protein, Hbs antigen  
N/Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid  
C/Species: hepatitis B virus, HBV  
A/Variety: subtype ayw4  
C/Date: 23-Nov-1994 #sequence\_revision 12-May-1995 #text\_change 26-Aug-1999  
C/Accession: S47407; JQ2071  
R/Plucienniczak, A.  
submitted to the EMBL Data Library, August 1994  
A/Description: Molecular cloning and sequencing of two complete genomes of polish isolat  
A/Reference number: S47404  
A/Accession: S47407  
A/Molecule type: DNA  
A/Residues: 1-389 <PLU>  
A/Cross-references: EMBL:Z35716; NID:g527435; PIDN:CAA84788.1; PID:g527439  
A/Experimental source: subtype ayw4  
R/Norder, H.; Hammes, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,  
J. Gen. Virol. 74, 1341-1348, 1993  
A/Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin  
A/Reference number: JQ2044; MUID:93329382; PMID:8336122  
A/Accession: JQ2071  
A/Molecule type: DNA  
A/Residues: 164-389 <NOR>  
A/Experimental source: subtype ayw3, strain 8950/90  
C/Genetics:  
A/Introns: 111/3  
C/Superfamily: hepatitis B virus surface antigen  
C/Keywords: surface antigen  
F/1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>  
F/1-108/Domain: pre-S1 domain #status predicted <PRE1>  
F/109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <F  
F/109-163/Domain: pre-S2 domain #status predicted <PRE2>

F/164-389/Product: surface antigen S (small envelope protein) #status predicted <SAG>

Query Match 96.7%; Score 59; DB 2; Length 389;  
Best Local Similarity 91.7%; Pred. No. 0.0029;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRRIGDP 12  
DB 149 PLSISFRRIGDP 160

RESULT 8  
S32202  
large surface antigen - hepatitis B virus (subtype ayw, isolate patient C1005)  
N/Contains: major surface antigen; middle surface antigen  
C/Species: hepatitis B virus, HBV  
A/Variety: subtype ayw, isolate patient C1005  
C/Date: 20-Feb-1995 #sequence\_revision 06-Dec-1996 #text\_change 06-Dec-1996  
C/Accession: S32202  
R/Pfeiler-Adams, S.; Schlayer, M.J.; Peters, T.; Hettler, F.; Gerok, W.; Rasenack, J.  
submitted to the EMBL Data Library, March 1993  
A/Description: Identification and sequence analysis of hepatitis B virus DNA in immunolc  
A/Reference number: S32202  
A/Accession: S32202  
A/Molecule type: DNA  
A/Residues: 1-389 <PRE>  
A/Cross-references: EMBL:X72702  
A/Experimental source: subtype ayw, isolate patient C1005  
C/Genetics:  
A/Introns: 111/3  
A/Genes: pre-S1/pre-S2/S  
C/Superfamily: hepatitis B virus surface antigen  
C/Keywords: surface antigen  
F/1-389/Product: large surface antigen (gene pre-S1/pre-S2/S) #status predicted <DSL>  
F/109-389/Product: middle surface antigen (gene S) #status predicted <MSA>  
F/164-389/Product: major surface antigen (gene S) #status predicted <MSA>

Query Match 96.7%; Score 59; DB 2; Length 389;  
Best Local Similarity 91.7%; Pred. No. 0.0029;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRRIGDP 12  
DB 149 PLSISFRRIGDP 160

RESULT 9  
T13474  
large surface antigen - hepatitis B virus (isolate 29Y11HC)  
N/Alternate names: envelope protein, Hbs antigen  
N/Contains: major surface antigen; middle surface antigen  
C/Species: hepatitis B virus, HBV  
A/Variety: isolate 29Y11HC  
C/Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 21-Jul-2000  
C/Accession: T13474  
R/Takahashi, K.; Akshane, Y.; Hino, K.; Ohta, Y.; Mshiro, S.  
Arch. Virol. 143, 2313-2326, 1998  
A/Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcino  
A/Reference number: Z17684; MUID:99129050; PMID:9930189  
A/Accession: T13474  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-382 <TAK>  
A/Cross-references: EMBL:AB014388; NID:g3582381; PIDN:BA432936.1; PID:g3582384  
A/Experimental source: Japanese patient with hepatocellular carcinoma isolate 29Y11HC  
C/Genetics:  
A/Introns: 123/2  
C/Superfamily: hepatitis B virus surface antigen  
C/Keywords: glycoprotein; surface antigen

Query Match 88.5%; Score 54; DB 2; Length 382;  
Best Local Similarity 83.3%; Pred. No. 0.025;  
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12  
 ||:|||||||  
 Db 142 PLSSIFSRIGDP 153

## RESULT 10

113469  
 large surface antigen - hepatitis B virus (isolate 09D09HCC)  
 N:Alternate names: envelope protein  
 N:Contains: major surface antigen; middle surface antigen  
 C:Species: hepatitis B virus, HBV  
 A:Variety: isolate 09D09HCC  
 C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 21-Jul-2000  
 C/Accession: T13469  
 R:Nakhaishi, K.; Akahane, Y.; Hino, K.; Ohta, Y.; Mishiro, S.  
 Arch. Virol. 143, 2313-2326, 1998  
 A>Title: Hepatitis B virus genomic sequence in the circulation of hepatocellular carcin  
 A/Reference number: Z17684; MUID:99129050; PMID:9930189  
 A/Accession: T13469  
 A:Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: DNA  
 A/Residues: 1-384 <TAK>  
 A/Cross-references: EMBL:AB014368; NID:G3551304; PIDN:BA432864.1; PID:G3551308  
 A:Experimental source: Japanese patient with hepatocellular carcinoma isolate 09D09HCC  
 C:Genetics:  
 A:Gene: S  
 A:Introns: 123/2  
 C:Superfamily: hepatitis B virus surface antigen  
 C:Keywords: glycoprotein; surface antigen

Query Match 88.5%; Score 54; DB 2; Length 384;  
 Best Local Similarity 83.3%; Pred. No. 0.025;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12  
 ||:|||||||  
 Db 144 PLSSIFSRIGDP 155

RESULT 11  
 S20745  
 surface antigen - hepatitis B virus (subtype ayw, patient C)  
 N:Alternate names: HBs antigen  
 N:Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid  
 C:Species: hepatitis B virus, HBV  
 A:Variety: subtype ayw, patient C  
 C>Date: 20-Feb-1995 #sequence\_revision 01-Nov-1996 #text\_change 26-Aug-1999  
 C/Accession: S20745  
 R:Irai, M.B.; Mazzoleni, A.P.; Balestrieri, A.; Mellis, A.; Porru, A.  
 Submitted to the EMBL Data Library, March 1992  
 A>Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negati  
 A/Reference number: S20745  
 A/Accession: S20745  
 A:Molecule type: DNA  
 A/Residues: 1-389 <RAI>  
 A/Cross-references: EMBL:X65257; NID:G59429; PIDN:CAA46349.1; PID:G59430  
 A:Experimental source: subtype ayw, patient C  
 C:Genetics:  
 A:Gene: S  
 C:Superfamily: hepatitis B virus surface antigen  
 C:Keywords: surface antigen  
 F:1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>  
 F:1-108/Domains: pre-S1 domain #status predicted <PRE1>  
 F:109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <P  
 F:109-163/Domains: pre-S2 domain #status predicted <PRE2>  
 F:164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 88.5%; Score 54; DB 2; Length 389;  
 Best Local Similarity 100.0%; Pred. No. 0.025;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LSSIFSRIGDP 12

Db 150 LSSIFSRIGDP 160  
 ||:|||||||  
 ||:|||||||

## RESULT 12

SAVIA  
 large surface antigen - hepatitis B virus (subtype adr)  
 N:Contains: major surface antigen; middle surface antigen  
 C:Species: hepatitis B virus, HBV  
 C>Date: 14-Nov-1983 #sequence\_revision 14-Nov-1983 #text\_change 07-May-1999  
 C/Accession: A03705; S04566; J02107; PQ0608  
 R:Ono, Y.; Ono, H.; Sasada, R.; Igatahshi, K.; Sugino, Y.; Nishioka, K.  
 Nucleic Acids Res. 11, 1747-1757, 1983  
 A>Title: The complete nucleotide sequences of the cloned hepatitis B virus DNA; subtype  
 A/Reference number: A93460; MUID:83168919; PMID:6300776  
 A/Accession: A03705  
 A:Molecule type: DNA  
 A/Residues: 1-400 <ONO>  
 A/Cross-references: GB:V00867  
 R:Rho, H.M.; Kim, K.; Hyun, S.W.; Kim, Y.S.  
 Nucleic Acids Res. 17, 2124, 1989  
 A>Title: The nucleotide sequence and reading frames of a mutant hepatitis B virus subtyp  
 A/Reference number: S04566; MUID:89183619; PMID:2928116  
 A/Accession: S04566  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A/Residues: 1-50, 'R', 52-66, 'YP', 69-129, 'V', 131-142, 'P', 144-164, 'S', 166-176, 'S', 178-338, '  
 A/Cross-references: EMBL:X14193  
 R:Norde, H.; Hammas, B.; Lee, S.D.; Bile, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus,  
 J. Gen. Virol. 74, 1341-1348, 1993  
 A>Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin  
 A/Reference number: J02044; MUID:93329382; PMID:8336122  
 A/Accession: J02107  
 A:Molecule type: DNA  
 A/Residues: 175-400 <NOR>  
 A:Experimental source: genogroup C, subtype adr, strain pBRHBadr4  
 R:Norde, H.; Courouce, A.M.; Magnus, L.O.  
 J. Gen. Virol. 73, 3141-3145, 1992  
 A>Title: Molecular basis of hepatitis B virus serotype variations within the four major  
 A/Reference number: PQ0453; MUID:93107848; PMID:1469553  
 A/Accession: PQ0608  
 A:Molecule type: DNA  
 A/Residues: 275-354 <NO2>  
 A:Experimental source: subtype adrq+, Bau

C:Genetics:  
 A:Gene: pre-S1/pre-S2/S  
 C:Superfamily: hepatitis B virus surface antigen  
 C:Keywords: glycoprotein; surface antigen  
 F:120-400/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>  
 F:175-400/Product: major surface antigen (gene S) #status predicted <MSA>  
 F:15,123,177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 88.5%; Score 54; DB 1; Length 400;  
 Best Local Similarity 83.3%; Pred. No. 0.026;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12  
 ||:|||||||  
 Db 160 PLSSIFSRIGDP 171

## RESULT 13

S35528  
 surface antigen - hepatitis B virus (subtype adr)  
 N:Alternate names: envelope protein; HBs antigen  
 N:Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid  
 C:Species: hepatitis B virus, HBV  
 A:Variety: subtype adr  
 C>Date: 09-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Jun-2000  
 C/Accession: S35528  
 R:Mukaido, M.; Kumazawa, T.; Hoshi, A.; Kawaguchi, R.; Hikiiji, K.  
 Nucleic Acids Res. 20, 6105, 1992  
 A>Title: The complete nucleotide sequence of hepatitis B virus, subtype adr (SRADR) and



```

A:Reference number: S35527; MUID:93096607; PMID:1461746
A:Accession: S35528
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-400 <MWK>
A:Cross-references: EMBL:DJ2980; NID:G221500; PIDN:BA02359.1; PID:G221502
A:Experimental source: subtype adr
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992
C:Genetics:
A:Gene: S
A:Introns: 165/2
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: surface antigen
F:1-400/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F:1-119/Domains: pre-S1 domain #status predicted <PRE1>
F:120-400/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <R
F:120-174/Domains: pre-S2 domain #status predicted <PRE2>
F:175-400/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Qy 1 PLSIFSRIGDP 12
Db 160 PLSIFSRIGDP 171

Query Match 88.5%; Score 54; DB 2; Length 400;
Best Local Similarity 83.3%; Pred. No. 0.026;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

RESULT 14
543492
Surface antigen - hepatitis B virus (subtype adr)
N:Alternate names: Hbs antigen
N:Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C:Species: hepatitis B virus, HBV
A:Variety: subtype adr
C:Date: 07-Sep-1994 #sequence revision 10-Nov-1995 #text_change 26-Aug-1999
A:Accession: S43492; PNO601; PNO602
R:Loncarevic, I. F.; Zentgraf, H.; Schroeder, C. H.
N:Clonacid Res. 18, 4940, 1990
A:Title: Sequence of a replication competent hepatitis B virus genome with a preX open r
A:Reference number: S12598; MUID:90370503; PMID:12395664
A:Accession: S43492
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-445 <LON>
A:Cross-references: EMBL:X52939; NID:G457780; PIDN:CAA37114.1; PID:G457783
R:MiMms, L.T.; Solomon, L.R.; Ebert, J.W.; Fields, H.
Biochem. Biophys. Res. Commun. 195, 186-191, 1993
A:Title: Unique pres sequence in a gibbon-derived hepatitis B virus variant.
A:Reference number: PNO601; MUID:93371402; PMID:8363598
A:Accession: PNO601
A:Molecule type: protein
A:Residues: 57-95 <MIM>
A:Accession: PNO602
A:Molecule type: protein
A:Residues: 165-196 <MI2>
C:Genetics:
A:Gene: S
A:Introns: 210/3
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: surface antigen
F:1-445/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F:1-144/Domains: pre-S1 domain #status predicted <PRE1>
F:165-445/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <R
F:165-219/Domains: pre-S2 domain #status predicted <PRE2>
F:220-445/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Qy 1 PLSIFSRIGDP 12
Db 160 PLSIFSRIGDP 171

Query Match 88.5%; Score 54; DB 2; Length 445;
Best Local Similarity 83.3%; Pred. No. 0.029;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 PLSIFSRIGDP 12

```

```

Db      : ||| |||| |
205    PISIFSRIGDP 216

RESULT 15
S67506
large surface antigen - hepatitis B virus (subtype ayw3, isolate Hope CH1357)
N:Alternate names: envelope protein; HBs antigen
N:Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C:Species: hepatitis B virus, HBV
A:Variety: subtype ayw3, isolate Hope CH1357
C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 26-Aug-1999
C:Accession: S67506
R:Norder: H.; Ebert, J.W.; Fields, H.A.; Mushatwar, I.K.; Magnus, L.O.
Virology 218, 214-223, 1996
A>Title: Complete sequencing of a gibbon hepatitis B virus genome reveals a unique genot
A:Reference number: S67503; MUID:96207410; PMID:8615024
A:Accession: S67506
A:Molecule type: DNA
A:Residues: 1-389 <NR>
A:Cross-references: EMBL:U46935; NID:g1814218; PIDN:AAB41952.1; PID:g1814222
A:Experimental source: subtype ayw3, isolate Hope CH1357
C:Genetics:
A:Gene: pre-S1/pre-S2/S
A:introns: 112/2
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: glycoprotein; surface antigen
F:1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F:1-108/Domin: pre-S1-specific domain #status predicted <PRE1>
F:109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <F
F:109-163/Domin: pre-S2-specific domain #status predicted <PRE2>
F:164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>
F:4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match          77.0%; Score 47; DB 2; Length 389;
Best Local Similarity 81.8%; Pred. No. 0.54;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      2 ISSIFSRIGDP 12
       : ||||| |||
Db      150 ISSIFSRIGDP 160

RESULT 16
S41870
surface antigen - hepatitis B virus (subtype ayw, patient M) (fragment)
N:Alternate names: envelope protein; HBs antigen
N:Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid
C:Species: hepatitis B virus, HBV
A:Variety: subtype ayw, patient M
C:Date: 06-Jan-1995 #sequence_revision 01-Nov-1996 #text_change 26-Aug-1999
C:Accession: S41870
R: Lai, M.E.; Mazzoleni, A.P.; Balestrieri, A.
submitted to the EMBL Data Library, January 1994
A:Description: Sequence analysis of HBV genomes isolated from patients with HBsAg chroma
A:Reference number: S41869
A:Accession: S41870
A:Molecule type: DNA
A:Residues: 1-378 <LA1>
A:Cross-references: EMBL:X77308; NID:g452613; PIDN:CAA54514.1; PID:g452614
A:Experimental source: subtype ayw, patient M
C:Genetics:
A:Gene: S
C:Superfamily: hepatitis B virus surface antigen
C:Keywords: surface antigen
F:1-378/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>
F:1-108/Domin: pre-S1 domain #status predicted <PRE1>
F:109-378/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <F
F:109-163/Domin: pre-S2 domain #status predicted <PRE2>
F:164-378/Product: surface antigen S (small envelope protein) #status predicted <PSD>

```

Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 PLSISFRRIGDP 12  
|||||:||||  
Db 149 PLSISARTGDP 160

RESULT 17  
S20753  
surface antigen - hepatitis B virus (subtype ayw, patient E)  
N/Alternate names: envelope protein; HBs antigen  
N/Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid  
C/Species: hepatitis B virus, HBV  
A/Variety: subtype ayw, patient E  
C/Date: 20-Feb-1995 #sequence\_revision 01-Nov-1996 #text\_change 26-Aug-1999  
C/Accession: S20753  
R/Jai, M.E.; Mazzoleni, A.P.; Balestrieri, A.; Melis, A.; Porru, A.  
submitted to the EMBL Data Library, March 1992  
A/Description: Sequence analysis of HBV genomes isolated from patients with HBsAg negativ  
A/Reference number: S20745  
A/Accession: S20753  
A/Molecule type: DNA  
A/Residues: 1-389 <LAI>  
A/Cross-references: EMBL:X65259; NID:g59439; PIDN:CAA46357.1; PID:g59440  
A/Experimental source: subtype ayw, patient E  
C/Genetics:

A/Gene: S  
C/Superfamily: hepatitis B virus surface antigen  
C/Keywords: surface antigen  
F/1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>  
F/1-108/Domain: pre-S1 domain #status predicted <PRE1>  
F/109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <F  
F/109-163/Domain: pre-S2 domain #status predicted <PRE2>  
F/164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 73.8%; Score 45; DB 2; Length 389;  
Best Local Similarity 75.0%; Pred. No. 1.3;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 PLSISFRRIGDP 12  
|||||:||||  
Db 149 PLSISARTGDP 160

RESULT 18  
B81194  
conserved hypothetical protein NMB0471 [imported] - Neisseria meningitidis (strain MC58  
C/Species: Neisseria meningitidis  
C/Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C/Accession: B81194  
R/Tetcelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A  
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;  
Li, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.  
Science 267, 1809-1815, 2000  
A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve  
A/Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A/Reference number: A81000; MUID:2015755; PMID:10710307  
A/Accession: B81194  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-358 <TET>  
A/Cross-references: GB:AE002404; GB:AE002098; NID:g7225697; PIDN:AAF40908.1; PID:g722569  
A/Experimental source: serogroup B, strain MC58  
C/Genetics:

A/Gene: NMB0471  
Query Match 70.5%; Score 43; DB 2; Length 358;  
Best Local Similarity 58.3%; Pred. No. 2.8;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 PLSISFRRIGDP 12  
|||:||||  
Db 121 PLGPVFRUGDP 132

RESULT 19  
A81831  
hypothetical protein NMA2014 [imported] - Neisseria meningitidis (strain Z2491 serogroup  
C/Species: Neisseria meningitidis  
C/Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C/Accession: A81831  
R/Parhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
Nature 404, 502-506, 2000  
A/Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
A/Reference number: A81775; MUID:2022556; PMID:10761919  
A/Accession: A81831  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1358 <PAR>  
A/Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85233.1; PID:g738064  
A/Experimental source: serogroup A, strain Z2491  
C/Genetics:

A/Gene: NMA2014  
Query Match 70.5%; Score 43; DB 2; Length 358;  
Best Local Similarity 58.3%; Pred. No. 2.8;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 PLSISFRRIGDP 12  
|||:||||  
Db 121 PLGPVFRUGDP 132

RESULT 20  
C87709  
ubiquitome biosynthesis protein Aarf, probable [imported] - Caulobacter crescentus  
C/Species: Caulobacter crescentus  
C/Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C/Accession: C87709  
R/Merham, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J  
B.; Laub, M.T.; DebRoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolor  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A/Title: Complete Genome Sequence of Caulobacter crescentus.  
A/Reference number: A87249; MUID:21173698; PMID:11259647  
A/Accession: C87709  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-502 <STO>  
A/Cross-references: GB:AE005673; NID:g13425475; PIDN:AAK25671.1; GSPDB:GN00148  
C/Genetics:

A/Gene: CC3709  
C/Superfamily: Synecocystis ABC transporter slr1919  
Query Match 67.2%; Score 41; DB 2; Length 502;  
Best Local Similarity 58.3%; Pred. No. 9.9;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 PLSISFRRIGDP 12  
|||:||||  
Db 119 PLDEIFAEIGEP 130

RESULT 21  
SAVICP  
large surface antigen - hepatitis B virus  
N/Contains: major surface antigen; middle surface antigen  
C/Species: hepatitis B virus, HBV  
C/Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 16-Jun-2000  
C/Accession: C28885; J02123  
J/Vaudin, M.; Wolstenholme, A.J.; Tsiquaye, K.N.; Zuckerman, A.J.; Harrison, T.J.  
J. Gen. Virol. 69, 1383-1389, 1988  
A/Title: The complete nucleotide sequence of the genome of a hepatitis B virus isolated  
A/Reference number: A97796; MUID:88258473; PMID:2838576  
A/Accession: C28885

A:Molecule type: DNA  
A:Residues: 1-389 <VAV>  
A:Cross-References: GB:D00220; NID:9221505; PIDN:BA00159.1; PID:9221508  
A:Experimental source: strain LSH, chimpanzee  
A>Note: the authors translated the codon GAG for residue 327 as Gln  
R:Order: H.; Hammes, B.; Lee, S.D.; Bille, K.; Courouce, A.M.; Mushahwar, I.K.; Magnus, J. Gen. Virol. 74, 1341-1348, 1993  
A>Title: Genetic relatedness of hepatitis B viral strains of diverse geographical origin  
A:Reference number: JQ2044; MUID:93329382; PMID:8336122  
A:Accession: JQ2123  
A:Molecule type: DNA  
A:Residues: 164-389 <NOR>  
A:Experimental source: subtype adw2, strain adwLSH  
C:Genetics:  
A:Gene: pre-S1/pre-S2/S; S  
C:Superfamily: hepatitis B virus surface antigen  
C:Keywords: glycoprotein; surface antigen  
F:109-389/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>  
F:164-389/Product: major surface antigen (gene S) #status predicted <MSA>  
F:4,112,166/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 65.6%; Score 40; DB 1; Length 389;  
Best Local Similarity 63.6%; Pred. No. 12;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 ISSIFSRIGDP 12  
DB 150 ISSIFSTRIGDP 160

RESULT 22  
S41871  
surface antigen - hepatitis B virus (subtype ayw, patient P)  
N:Alternate names: envelope protein, HBs antigen  
N:Contains: surface antigen pre-S1 (large envelope protein); surface antigen pre-S2 (mid  
C:Species: hepatitis B virus, HBV  
A:Variety: subtype ayw, patient P  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 26-Aug-1999  
C:Accession: S41871  
R:Lat, M.E.; Marzotani, A.P.; Balestrieri, A.  
Submitted to the EMBL Data Library, January 1994  
A:Description: Sequence analysis of HBV genomes isolated from patients with HBsAg chroni  
A:Reference number: S41869  
A:Accession: S41871  
A:Molecule type: DNA  
A:Residues: 1-389 <LAT>  
A:Cross-References: EMBL:X77310; NID:9452615; PIDN:CA54516.1; PID:9452616  
A:Experimental source: subtype ayw, patient P  
C:Genetics:  
A:Gene: S  
C:Superfamily: hepatitis B virus surface antigen  
C:Keywords: surface antigen  
F:1-389/Product: surface antigen pre-S1 (large envelope protein) #status predicted <PS1>  
F:1-108/Domain: pre-S1 domain #status predicted <PRE1>  
F:109-389/Product: surface antigen pre-S2 (middle envelope protein) #status predicted <H  
F:164-389/Domain: pre-S2 domain #status predicted <PRE2>  
F:164-389/Product: surface antigen S (small envelope protein) #status predicted <PSD>

Query Match 65.6%; Score 40; DB 2; Length 389;  
Best Local Similarity 63.6%; Pred. No. 12;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 ISSIFSRIGDP 12  
DB 150 ISSIFSTRIGDP 160

RESULT 23  
AF2349  
hypothetical protein al14349 [imported] - Nostoc sp. (strain PCC 7120)  
C:Species: Nostoc sp. PCC 7120  
A>Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C:Accession: AF2349  
R:Kazuko, T.; Nakamura, Y.; Wolk, C.P.; Korfiz, T.; Sasamoto, S.; Watanabe, A.; Iritschi  
Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 6, 205-213, 2001  
A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AF2349  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-399 <KUR>  
A:Cross-References: GB:BA00019; PIDN:BA076048.1; PID:917133485; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: al14349

Query Match 63.9%; Score 39; DB 2; Length 399;  
Best Local Similarity 58.3%; Pred. No. 18;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12  
DB 351 PMSIFSTRIGDP 362

RESULT 24  
S42226  
hypothetical large surface antigen/mevalonate kinase (EC 2.7.1.36) mutant fusion protein  
C:Species: hepatitis B virus, HBV  
C:Date: 13-Jan-1995 #sequence\_revision 17-Mar-1997 #text\_change 20-Oct-2000  
C:Accession: S42226  
R:Graef, E.; Caselmann, W.H.; Wells, J.; Koshy, R.  
Oncogene 9, 81-87, 1994  
A>Title: Insertional activation of mevalonate kinase by hepatitis B virus DNA in a human  
A:Reference number: S42226; MUID:94134441; PMID:8302606  
A:Accession: S42226  
A:Molecule type: mRNA  
A:Residues: 1-557 <GRA>  
A:Cross-References: EMBL:X75311; NID:9450345; PIDN:CA53059.1; PID:9450346  
C:Keywords: phosphotransferase  
F:1-157/Region: hepatitis B virus large surface antigen (fragment)  
F:162-557/Region: human mevalonate kinase (fragment)

Query Match 63.9%; Score 39; DB 4; Length 557;  
Best Local Similarity 72.7%; Pred. No. 26;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 ISSIFSRIGDP 12  
DB 42 ISSIFSTRIGDP 52

RESULT 25  
AF2895  
hypothetical protein Atu2599 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
C:Accession: AF2895  
R:Wood, D.W.; Setudal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavyn, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krepsan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A>Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AF2895  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-278 <KUR>  
A:Cross-References: GB:AE00688; PIDN:AAU43580.1; PID:917741096; GSPDB:GN00186  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu2599

A:Map position: circular chromosome

Query Match 62.3%; Score 38; DB 2; Length 278;  
Best Local Similarity 63.6%; Pred. No. 19;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 PLSISFRIGDP 11  
|||:||||  
DB 211 PLUKFSRFGD 221

#### RESULT 26

A:86807

hypothetical protein yoiB [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001

C/Accession: A86807

R:Okamoto, H.; Tenda, F.; Sakugawa, H.; Saetsoewignjo, R.I.; Imai, M.; Miyakawa, Y.; Ehrlich, J. Gen. Virol. 69, 2575-2583, 1988

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss

F/4,112/Binding site: carbohydrate (Asn) (covalent) #status predicted

A:Reference number: J50253; MUID:89010694; PMID:11337471

A:Accession: A86807

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-315 <STO>

A:Cross-references: GB:AE005176; PID:g12724449; PIDN:AAK0555.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: yoiB

Query Match 62.3%; Score 38; DB 2; Length 315;  
Best Local Similarity 70.0%; Pred. No. 22;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFRIG 10  
|||||:  
DB 21 PLSISYNOIG 30

#### RESULT 27

SAVLJ1

large surface antigen - hepatitis B virus (subtype adw, strain Japan/BJDW233)

N:Contains: major surface antigen; middle surface antigen

C:Species: hepatitis B virus, HBV

C>Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 14-Nov-1997

C/Accession: G28925

R:Okamoto, H.; Tenda, F.; Sakugawa, H.; Saetsoewignjo, R.I.; Imai, M.; Miyakawa, Y.; Ehrlich, J. Gen. Virol. 69, 2575-2583, 1988

A:Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of surf

A:Reference number: J50253; MUID:89010694; PMID:1171552

A:Accession: G28925

A:Molecule type: DNA

A:Residues: 1-389 <OKA>

A:Cross-references: GB:D00329; NID:g221497

C:Genetics:

A:Gene: pre-S1/pre-S2/S

C:Superfamily: hepatitis B virus surface antigen

C:Keywords: glycoprotein; surface antigen (gene pre-S2/S) #status predicted <DSA>

F/109-389/Product: middle surface antigen (gene S) #status predicted <MSA>

F/175-389/Product: major surface antigen (gene S) #status predicted <MSA>

F/4,112/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 62.3%; Score 38; DB 1; Length 389;  
Best Local Similarity 63.6%; Pred. No. 28;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LSSISFRIGDP 12  
|||:||||  
DB 150 ISSILSKTGP 160

#### RESULT 28

SAVLJ2

large surface antigen - hepatitis B virus (subtype adw, strain Okinawa/PODW282)

N:Contains: major surface antigen; middle surface antigen

C:Species: hepatitis B virus, HBV

C>Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 14-Nov-1997

C/Accession: H28925

R:Okamoto, H.; Tenda, F.; Sakugawa, H.; Saetsoewignjo, R.I.; Imai, M.; Miyakawa, Y.; Ehrlich, J. Gen. Virol. 69, 2575-2583, 1988

A:Title: Typing hepatitis B virus by homology in nucleotide sequence: comparison of surf

A:Reference number: J50253; MUID:89010694; PMID:1171552

A:Accession: H28925

A:Molecule type: DNA

A:Residues: 1-389 <OKA>

A:Cross-references: GB:D00330; NID:g221498

C:Genetics:

A:Gene: pre-S1/pre-S2/S

C:Superfamily: hepatitis B virus surface antigen

C:Keywords: glycoprotein; surface antigen (gene pre-S2/S) #status predicted <DSA>

F/109-389/Product: middle surface antigen (gene S) #status predicted <MSA>

F/164-389/Product: major surface antigen (gene S) #status predicted <MSA>

F/4,112/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 62.3%; Score 38; DB 1; Length 389;  
Best Local Similarity 63.6%; Pred. No. 28;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

#### RESULT 29

F86852

amino acid permease ysjA [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis

C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001

C/Accession: F86852

R:Okamoto, H.; Tenda, F.; Sakugawa, H.; Saetsoewignjo, R.I.; Imai, M.; Miyakawa, Y.; Ehrlich, J. Gen. Virol. 69, 2575-2583, 1988

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss

A:Reference number: A86825; MUID:21235186; PMID:11337471

A:Accession: F86852

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-449 <STO>

A:Cross-references: GB:AE005176; PID:g12724850; PIDN:AAK05920.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: ysjA

C:Superfamily: arginine permease

Query Match 62.3%; Score 38; DB 2; Length 449;  
Best Local Similarity 50.0%; Pred. No. 32;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 PLSISFRIGDP 12  
|||:||||  
DB 276 PFATIFARVGP 287

#### RESULT 30

S67434

probable sterol O-acetyltransferase (EC 2.3.1.26) SPAC13G7.05 [similarity] - fission yeast

C:Species: Schizosaccharomyces pombe

C>Date: 20-Jul-1996 #sequence\_revision 13-Mar-1997 #text\_change 08-Sep-2000

C/Accession: S67434; T57655

R:Connor, R.; Churcher, C.M. submitted to the EMBL Data Library, February 1996

A:Reference number: S67430

A:Accession: S67434

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-537 <CON>

A:Cross-references: EMBL:Z69729; NID:g1204167; PIDN:CAA93593.1; PID:g223696; PID:g120417

RESULT 32  
 H71527  
 C:Probable excinuclease ABC chain A - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
 C:Species: Chlamydia trachomatis  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: H71527  
 R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marache, R.; Aravind, L.; Mitchell,  
 Science 282, 754-759, 1998  
 A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tra  
 A:Reference number: A71570; MUID:99000809; PMID:9784136  
 A:Accession: H71527  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1786 <ABN>  
 A:Cross-references: GB:AE001306; GB:AE001273; NID:G3328748; PID:G3328752  
 A:Experimental source: serotype D, strain UW-3/Cx  
 C:Genetics:  
 A:Gene: uvrA  
 C:Superfamily: Chlamydia trachomatis probable excinuclease ABC chain A; ATP-binding case

RESULT 34  
882477  
hypothetical protein VCA0303 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)  
C:Species: *Vibrio cholerae*  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: 882477  
R:Heldelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;  
charadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Baer, S.; Qin, H.; Dragol, I.; Sellers, F.  
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
Reference number: A82035; MUID:20406833; PMID:10952301

|                         |       |              |          |            |
|-------------------------|-------|--------------|----------|------------|
| Query Match             | 60.7% | Score 37     | DB 2     | Length 134 |
| Best Local Similarity   | 72.7% | Pred. NO. 13 |          |            |
| Matches 8; Conservative | 1     | Mismatches 2 | Indels 0 | Gaps 0     |

QY 2 LSSIFSRIGDP 12  
|||||: ||

Db 74 LSSIFSGWGP 84

RESULT 35

A:Accession: A97370

A:Title: hypothetical protein AGR\_C\_144 [imported] - Agrobacterium tumefaciens (strain C58, Cere

C:Species: Agrobacterium tumefaciens

C>Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002

C/Accession: A97370

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Marwitz, B.;

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; PMID:21608551; PMID:11743194

A:Accession: A97370

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-229 <KUR>

A:Cross-references: GB:AE007869; PIDN:AAK85914.1; PID:G15154963; GSPDB:GN00169

C:Genetics:

A:Gene: AGR\_C\_144

A:Map position: circular chromosome

Query Match 60.7%; Score 37; DB 2; Length 229;

Best Local Similarity 50.0%; Pred. No. 24;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12

Db 141 PCATLSARIGDP 152

RESULT 36

A:Accession: A12587

A:Title: conserved hypothetical protein Atu0094 [imported] - Agrobacterium tumefaciens (strain C5

C:Species: Agrobacterium tumefaciens

C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002

C/Accession: A12587

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monke, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClell

, Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:21608550; PMID:11743193

A:Accession: A12587

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-229 <KUR>

A:Cross-references: GB:AE008688; PIDN:AAU41119.1; PID:G17738412; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: Atu0094

A:Map position: circular chromosome

Query Match 60.7%; Score 37; DB 2; Length 229;

Best Local Similarity 50.0%; Pred. No. 24;

Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 PLSIFSRIGDP 12

Db 141 PCATLSARIGDP 152

RESULT 37

A:Accession: A90956

A:Title: probable integrase [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952

C:Species: Escherichia coli

C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001

C/Accession: A90956

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gawaara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc

A:Reference number: A99629; PMID:21156231; PMID:11258796

A:Accession: A90956

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-333 <HAV>

A:Cross-references: GB:BA000007; PIDN:BA936040.1; PID:G13362085; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: Ece2617

C:Superfamily: phage P22 integrase

Query Match 60.7%; Score 37; DB 2; Length 333;

Best Local Similarity 63.6%; Pred. No. 36;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LSSIFSRIGDP 12

Db 84 LRAICSRIGDP 94

RESULT 38

F85804

Integrase for prophage CP-933r [imported] - Escherichia coli (strain O157:H7, substrain

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C/Accession: F85804

R:Peria, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glaesner, J.D.; Rose, D.J.; Mayhe

iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamoussis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: AB5480; PMID:21074935; PMID:11206551

A:Accession: F85804

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-333 <STO>

A:Cross-references: GB:AE005174; NID:G12515975; PIDN:AA656898.1; GSPDB:GN00145; UWGP:225

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: intC

C:Superfamily: phage P22 integrase

Query Match 60.7%; Score 37; DB 2; Length 333;

Best Local Similarity 63.6%; Pred. No. 36;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LSSIFSRIGDP 12

Db 84 LRAICSRIGDP 94

RESULT 39

I50516

retinoid X receptor delta - zebra fish

C:Species: Brachydanio rerio (zebra fish)

C>Date: 13-Mar-1997 #sequence\_revision 13-Mar-1997 #text\_change 02-Aug-2002

C/Accession: I50516

R:Jones, B.B.; Ohno, C.K.; Allenby, G.; Boffa, M.B.; Levin, A.A.; Grippo, J.F.; Petkovic

Mol. Cell. Biol. 15, 5226-5234, 1995

A:Title: New retinoid X receptor subtypes in zebra fish (Danio rerio) differentially mod

A:Reference number: A57301; PMID:96009547; PMID:7565671

A:Accession: I50516

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-422 <JON>

A:Cross-references: EMBL:U29941; NID:G1046296; PIDN:AA659721.1; PID:G1046297

C:Genetics:

A:Gene: RXR

C:Superfamily: retinoic acid receptor alpha; erba transforming protein homology

C:Keywords: zinc finger

F:88-341/Domain: erba transforming protein homology <ERBA>

Query Match 60.7%; Score 37; DB 2; Length 422;  
 Best Local Similarity 58.3%; Pred. No. 47;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGDP 12  
 |||::|||  
 Db 21 PLSAVSSSIGSP 32

## RESULT 40

T21115

hypothetical protein F19C6.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C:Accession: T21115

R:Haris, B.

submitted to the EMBL Data Library, January 1995

A:Reference number: Z19376

A:Accession: T21115

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-495 &lt;NII&gt;

A:Cross-references: EMBL:Z48006; PIDN:CA88050.1; GSPDB:GN00028; CESP:F19C6.2

A:Experimental source: clone F19C6

C:Genetics:

A:Gene: CESP:F19C6.2

A:Map position: X

A:Introns: 44/1; 117/3; 153/3; 213/3; 238/3; 334/1; 387/2; 417/3

C:Superfamily: Caenorhabditis elegans hypothetical protein F19C6.2

Query Match 60.7%; Score 37; DB 2; Length 495;  
 Best Local Similarity 45.5%; Pred. No. 56;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 PLSSIFSRIGD 11  
 |::|::|  
 Db 278 PVSGLFQKVGD 288

Search completed: November 6, 2003, 15:13:16  
 UOD time : 23 secs

**THIS PAGE BLANK (USPTO)**





FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 75 75 O -> E (IN STRAIN LATVIA).  
 FT VARIANT 147 147 A -> S (IN STRAIN LATVIA).  
 FT VARIANT 150 150 L -> I (IN STRAIN LATVIA).  
 FT VARIANT 288 290 MTT -> TTP (IN STRAIN LATVIA).  
 SQ SEQUENCE 389 AA; 42766 MW; 6DC9E862DA694F63 CRC64;

Query Match 100.0%; Score 61; DB 1; Length 389;  
 Best Local Similarity 100.0%; Pred. No. 0.00054;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFIRIGDP 12  
 DB 149 PLSISFIRIGDP 160

## RESULT 2

VSMA\_HPBVA STANDARD; PRT; 389 AA.

AC P24035;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 01-MAR-1992 (Rel. 21, Last annotation update)  
 DE Major surface antigen precursor.  
 GN S.  
 OS Hepatitis B virus (strain alpha1).  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OX NCBI\_TaxID=10411;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90266476; PubMed=2345966;  
 RA Tong S., Li J., Vitvitski L., Trepo C.;  
 RT "Active hepatitis B virus replication in the presence of anti-HBe is  
 associated with viral variants containing an inactive pre-C region.";  
 RL Virology 176:596-603(1990).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
 DR EMBL\_M32138; AAA45502.1; .  
 DR PIR; B34773; SAVLA1.  
 DR InterPro; IPR000349; Hepvir\_surfAg.  
 DR Pfam; PF00695; VSMA; 1.  
 KM Antigen.  
 FT PROPEP 1 163  
 FT CHAIN 164 389  
 FT CARBOHYD 4 4 MAJOR SURFACE ANTIGEN.  
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 389 AA; 42733 MW; B06D0D42B23B5 CRC64;

Query Match 96.7%; Score 59; DB 1; Length 389;  
 Best Local Similarity 91.7%; Pred. No. 0.0013;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFIRIGDP 12  
 DB 149 PLSISFIRIGDP 160

## RESULT 3

VSMA\_HPBVA STANDARD; PRT; 389 AA.

AC P03139;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 01-OCT-1989 (Rel. 12, Last annotation update)  
 DE Major surface antigen precursor.

GN S.  
 OS Hepatitis B virus (subtype adw).  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OX NCBI\_TaxID=10411;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=81012115; PubMed=399329;  
 RA Pasek M., Goto T., Gilbert W., Zink B., Schaller H., Mackay P.,  
 RA Leadbetter G., Murray K.;  
 RT "Hepatitis B virus genes and their expression in E. coli.";  
 RL Nature 282:575-579(1979).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC  
 DR EMBL\_J02202; AAA45487.1; ALT\_INIT.  
 DR PIR; A93217; SAVLAJ.  
 DR InterPro; IPR000349; Hepvir\_surfAg.  
 DR Pfam; PF00695; VSMA; 1.  
 KM Antigen.  
 FT PROPEP 1 163  
 FT CHAIN 164 389  
 FT CARBOHYD 4 4 MAJOR SURFACE ANTIGEN.  
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 389 AA; 42801 MW; 1069FE8B4679A669 CRC64;

Query Match 96.7%; Score 59; DB 1; Length 389;  
 Best Local Similarity 91.7%; Pred. No. 0.0013;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFIRIGDP 12  
 DB 149 PLSISFIRIGDP 160

## RESULT 4

VSMA\_HPBVA STANDARD; PRT; 400 AA.

AC P03140;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Major surface antigen precursor.

GN S.  
 OS Hepatitis B virus (subtype adr).  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OX NCBI\_TaxID=106820;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=83168919; PubMed=6300776;  
 RA Ono Y., Ono H., Sasada R., Igataashi K., Sugino Y., Nishiohka K.;  
 RT "The complete nucleotide sequences of the cloned hepatitis B virus  
 DNA; subtype adr and adw.";  
 RL Nucleic Acids Res. 11:1747-1757(1983).

RN [2]  
 RP SEQUENCE OF 176-350 FROM N.A.  
 RX MEDLINE=85107103; PubMed=3968537;  
 RA Takeshima H., Inokoshi J., Namiki M., Shimada J., Omura S.;  
 RT "Structural analysis of the gene coding for hepatitis B virus surface  
 antigen and its product.";  
 RL J. Gen. Virol. 66:193-200(1985).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; V00867; CAA24234.1; ALT\_INIT.  
 DR PIR; A03705; SAVLA.  
 DR InterPro; IPR000349; Hepvir\_surfag.  
 DR Pfam; PF00695; VMSA; 1.

KW Antigen.  
 FT PROPEP 1 174  
 FT CHAIN 175 400  
 FT CARBOHYD 15 15  
 FT CARBOHYD 123 123  
 FT CARBOHYD 177 177  
 FT CARBOHYD 320 320  
 SQ SEQUENCE 400 AA; 43685 MW; 16DB0B15C6B97A CRC64;

Query Match  
 Best Local Similarity 83.3%; Score 54; DB 1; Length 400;  
 Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PSSISFRIGDP 12  
 Db 160 PSSISFRIGDP 171

RESULT 5  
 VMSA\_HPBV4  
 ID VMSA\_HPBV4 STANDARD; PRT; 400 AA.  
 AC P12934; O67860;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Major surface antigen precursor.  
 DE S.  
 OS Hepatitis B virus (subtype adr4).  
 CC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OK NCBI\_TaxID=10409;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=83246570; PubMed=6306594;  
 RA Fujiyama A., Miyahara A., Nozaki C., Yoneyama T., Ohomo N.,  
 RA Matubara K.,  
 RA "Cloning and structural analyses of hepatitis B virus DNAs, subtype  
 RT adr.";  
 RT Nucleic Acids Res. 11:4601-4610(1983).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X01587; CAA25747.1; ALT\_INIT.  
 DR EMBL; X01587; CAA25743.1; ALT\_INIT.  
 DR InterPro; IPR000349; Hepvir\_surfag.  
 DR Pfam; PF00695; VMSA; 1.  
 KW Antigen.  
 FT PROPEP 1 174  
 FT CHAIN 175 400  
 FT CARBOHYD 15 15  
 FT CARBOHYD 123 123  
 FT CARBOHYD 177 177  
 FT CARBOHYD 320 320  
 SQ SEQUENCE 400 AA; 43542 MW; 01053BBD24053F9 CRC64;

Query Match  
 Best Local Similarity 75.4%; Score 46; DB 1; Length 400;  
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PSSISFRIGDP 12  
 Db 160 PSSISFRIGDP 171

RESULT 6  
 VMSA\_HPBVP  
 ID VMSA\_HPBVP STANDARD; PRT; 400 AA.  
 AC Q02317;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-JUL-1993 (Rel. 26, Last annotation update)  
 DE Major surface antigen precursor.  
 DE S.  
 OS Hepatitis B virus (subtype adr / strain Philippines/GFDM294).  
 CC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OK NCBI\_TaxID=31514;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Estacio R.C., Chavez C.C., Okamoto H., Lingao A.L., Reyes M.T.,  
 RA Domingo E., Mayumi M.,  
 RA "Nucleotide sequence of a hepatitis B virus genome of subtype adr  
 RT isolated from a Filipino: comparison with the reported three genomes  
 RT of the same subtype."  
 RT J. Gastroenterol. Hepatol. 3:215-222(1988).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M57663; AAA69680.1; -  
 DR InterPro; IPR000349; Hepvir\_surfag.  
 DR Pfam; PF00695; VMSA; 1.  
 KW Antigen.  
 FT PROPEP 1 174  
 FT CHAIN 175 400  
 FT CARBOHYD 15 15  
 FT CARBOHYD 123 123  
 FT CARBOHYD 177 177  
 FT CARBOHYD 320 320  
 SQ SEQUENCE 400 AA; 43740 MW; 7EBD53A004136BD2 CRC64;

Query Match  
 Best Local Similarity 72.1%; Score 44; DB 1; Length 400;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PSSISFRIGDP 12  
 Db 161 PSSISFRIGDP 171

RESULT 7  
 BD02\_HUMAN  
 ID BD02\_HUMAN STANDARD; PRT; 64 AA.  
 AC O15263;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Beta-defensin 2 precursor (hBD-2) (Skin-antimicrobial peptide 1)  
 DE (SAP1).  
 DE DEF84 OR DEF82 OR DEF8102.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OK NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP TISSUE=Skin;

RX MEDLINE=97345625; PubMed=9202117;  
 RA Harder J., Bartels J.H., Christophers E., Schroeder J.-M.;  
 RT "A peptide antibiotic from human skin.";  
 RL Nature 387:861-861(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=99051334; PubMed=9831658;  
 RA Liu L., Wang L., Jia H.P., Zhao C., Heng H.H.Q., Schutte B.C.,  
 RT McCray P.B. Jr., Ganz T.;  
 RL "Structure and mapping of the human beta-defensin HBD-2 gene and its  
 expression at sites of inflammation.";  
 RL Gene 222:237-244(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20072673; PubMed=10603376;  
 RA Diamond G., Kaiser V., Rhodes J., Russell J.P., Bevins C.L.;  
 RT "Transcriptional regulation of beta-defensin gene expression in  
 tracheal epithelial cells.";  
 RL Infect. Immun. 68:113-119(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Harder J., Schroeder J.M.;  
 RT "Transcriptional regulation of the human beta-defensin-2 (HBD-2).";  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SYNTHESIS OF 24-64.  
 RX MEDLINE=2007551; PubMed=12010514;  
 RA Kluever E., Schulz A., Forssmann W.-G., Adernann K.;  
 RT "Chemical synthesis of beta-defensins and LEAP-1/hepcidin.";  
 RL J. Pept. Res. 59:241-248(2002).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (1.35 ANGSTROMS).  
 RX MEDLINE=20490730; PubMed=10906336;  
 RA Hoover D.M., Rajashankar K.R., Blumenthal R., Puri A., Oppenheim J.J.,  
 RT Chertov O., Lubkowski J.;  
 RL "The structure of human beta-defensin-2 shows evidence of higher  
 order oligomerization.";  
 RL J. Biol. Chem. 275:32911-32918(2000).  
 RN [7]  
 RP STRUCTURE BY NMR OF 28-64.  
 RX MEDLINE=21571984; PubMed=11714914;  
 RA Bauer F., Schweimer K., Kluever E., Conejo-Garcia J.R.,  
 RT Forssmann W.-G., Rosch P., Adernann K., Sticht H.;  
 RL "Structure determination of human and murine beta-defensins reveals  
 structural conservation in the absence of significant sequence  
 similarity.";  
 RL Protein Sci. 10:2470-2479(2001).  
 CC -1- FUNCTION: HAS ANTIBACTERIAL ACTIVITY (POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN THE SKIN AND RESPIRATORY TRACT.  
 CC -1- INDUCTION: By inflammation.  
 CC -1- SIMILARITY: BELONGS TO THE BETA-DEFENSIN FAMILY. LAP/TAP  
 CC SUBFAMILY.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Z71389; CAA95992.1; -;  
 CC EMBL; AF040153; AAC3549.1; -;  
 CC EMBL; AF071216; AAC69554.1; -;  
 CC EMBL; AJ000152; CAB65126.1; -;  
 CC PDB; 1FD3; 01-NOV-00.  
 CC PDB; 1PD4; 01-NOV-00.  
 CC PDB; 1E40; 26-NOV-01.  
 CC PDB; 1F00; 31-DEC-02.  
 CC Genew; HGNC:2767; DEFB4.  
 CC MIM; 602215; -.

DR GO; GO:0006935; P:chemotaxis; TAS.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalling. . .; TAS.  
 DR GO; GO:0006955; P:immune response; TAS.  
 DR GO; GO:0006513; P:response to pest/pathogen/parasite; TAS.  
 DR InterPro; IPR001855; Defensin\_beta.  
 DR InterPro; IPR006080; Defensin\_mammal.  
 DR Pfam; PF00711; Defensin\_beta; 1.  
 DR SMART; SM00048; DEFSN; 1.  
 KM Antibiotic; Signal; 3D-structure.  
 FT SIGNAL 1 23  
 FT CHAIN 24 64  
 FT DISULFID 31 60  
 FT DISULFID 38 53  
 FT DISULFID 43 61  
 FT STRAND 26 26  
 FT HELIX 28 33  
 FT TURN 34 35  
 FT STRAND 37 39  
 FT TURN 45 46  
 FT STRAND 48 52  
 FT TURN 56 57  
 FT STRAND 59 62  
 SQ SEQUENCE 64 AA; 7038 MW; 05D6454CE7ACD10E CRC64;  
 Query Match 65.6%; Score 40; DB 1; Length 64;  
 Best Local Similarity 58.3%; Pred. No. 0.82;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 PLSTFSRIGDP 12  
 Db 17 PLPGVFGIGDP 28  
 RESULT 8  
 VMSA\_HPBVL STANDARD; PRT; 389 AA.  
 ID VMSA\_HPBVL  
 AC P12911;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 01-APR-1990 (Rel. 14, Last annotation update)  
 DE Major surface antigen precursor.  
 GN S.  
 OS Hepatitis B virus (strain 1sh / chimpanzee isolate).  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OX NCBI\_Taxid=10414;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88258473; PubMed=2838576;  
 RA Vaudin M., Wolstenholme A.J., Tsiquaye K.N., Zuckerman A.J.,  
 RT Harrison T.J.;  
 RL "The complete nucleotide sequence of the genome of a hepatitis B  
 virus isolated from a naturally infected chimpanzee.";  
 RL J. Gen. Virol. 69:1383-1389(1988).  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D00220; BAA00159.1; -;  
 DR PIR; C28885; SAVLCP.  
 DR InterPro; IPR000349; Hepvir\_surfAg.  
 DR Pfam; PF00695; VMSA; 1.  
 KM Antigen.  
 FT PROPEP 1 163  
 FT CHAIN 164 389  
 FT CARBOHYD 4 4  
 FT CARBOHYD 112 112  
 FT CARBOHYD 166 166  
 FT CARBOHYD 309 309  
 MAJOR SURFACE ANTIGEN.  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

SQ  SEQUENCE 369 AA; 42539 MW; 87D1P93362A10DBD CRC64;
      Query Match      65.6%; Score 40; DB 1; Length 389;
      Best Local Similarity 63.6%; Pred. No. 5.3;
      Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy  2  LSSIFSRIGDP 12
      :|||:||||
Db  150  ISSVFTTGDP 160

RESULT 9
VMSA_HPBVT
ID  _VMSA_HPBVT  STANDARD; PRT; 389 AA.
AC  P17398;
DT  01-AUG-1990 (Rel. 15, Created)
DT  01-AUG-1990 (Rel. 15, Last sequence update)
DT  01-AUG-1992 (Rel. 23, Last annotation update)
DE  Major surface antigen precursor.
GN  S.
OS  Hepatitis B virus (subtype adw / strain Japan/GJDM233).
OC  Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX  NCHI_TaxID=10413;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=89010694; PubMed=3171552;
RA  Okamoto H., Tsuda F., Sakugawa H., Sastrosewignjo R.I., Imai M.,
RA  Miyakawa Y., Mayumi M.;
RT  "Typing hepatitis B virus by homology in nucleotide sequence:
RL  comparison of surface antigen subtypes.",
RT  J. Gen. Virol. 63:2575-2583(1988).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on items
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb.ch/announce/
CC  or send an email to license@isb.ch).
CC  -----
DR  EMBL; D00329; -; NOT_ANNOTATED_CDS.
DR  PIR; G28925; SAVI.L1.
DR  InterPro; IPR000349; Hepvlt_surfAg.
DR  Pfam; PF00695; VMSA; 1.
DR  Antigen.
FT  KW PROPEP.
FT  FT PROPEP 1 163
FT  CHAIN 164 389 MAJOR SURFACE ANTIGEN.
FT  CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT  CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ  SEQUENCE 389 AA; 42603 MW; 72F8D481924DC104 CRC64;

Query Match 62.3%; Score 38; DB 1; Length 389;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Oy  2  LSSIFSRIGDP 12
      :|||:||||
Db  150  ISSVFTTGDP 160

RESULT 10
VMSA_HPBVO
ID  _VMSA_HPBVO  STANDARD; PRT; 389 AA.
AC  P17399;
DT  01-AUG-1990 (Rel. 15, Created)
DT  01-AUG-1990 (Rel. 15, Last sequence update)
DT  01-AUG-1992 (Rel. 23, Last annotation update)
DE  Major surface antigen precursor.
GN  S.
OS  Hepatitis B virus (subtype adw / strain Okinawa/GDDM282).
OC  Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

```

| Query | Match | Best local similarity | Score      | DB              | Length        |
|-------|-------|-----------------------|------------|-----------------|---------------|
| Qy    | 2     | LSIFSRGDP 12          | 62.3%      | DB 1            | 389           |
| Db    | 150   | ISSILSKGDP 160        | 63.6%      | Pred. No. 13    |               |
|       |       |                       | Matches 7; | Conservative 2; | Mismatches 2; |
|       |       |                       |            | Indels 0;       | Gaps 0;       |

RESULT 11

| ID | AREH_SCHPO  | STANDARD;                         | PRT; | 537 AA. |
|----|---|-----------------------------------|------|---------|
| AC | Q10269;   |                                   |      |         |
| DT | 01-OCT-1996   | (Rel. 34, Created)                |      |         |
| DT | 01-OCT-1996   | (Rel. 34, Last sequence update)   |      |         |
| DT | 28-FEB-2003   | (Rel. 41, Last annotation update) |      |         |
| DE | Probable sterol O-acetyltransferase (EC 2.3.1.26) (Sterol-ester synthase).  |                                   |      |         |
| GN | SPAC137.05.   |                                   |      |         |
| OS | Schizosaccharomyces pombe (Fission yeast).  |                                   |      |         |
| CC | Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  |                                   |      |         |
| CC | Schizosaccharomycetales; Schizosaccharomycetaceae;  |                                   |      |         |
| CC | Schizosaccharomycetes.  |                                   |      |         |
| OX | NCBI_TaxID=4896;  |                                   |      |         |
| RN | [1]   |                                   |      |         |
| RP | SEQUENCE FROM N.A.  |                                   |      |         |
| RC | STRAIN=972;   |                                   |      |         |
| FX | MEDLINE=21848401; PubMed=11859360;  |                                   |      |         |
| RA | Wood V., Williams M., Rajandream M.A., Lyne M., Lyne R., Stewart A., Sgouros V., Peat N., Hayles J., Baker S., Basham D., Bowman S., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares S., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Wolckele G., Aert R., Robben J., Grymptre B., Weltjens I., Vansteelt E., Rieger M., Schaefer M., Mueller-Auer S., |                                   |      |         |

RA Gabel C., Fuchs M., Filtz C., Holzer E., Moesli D., Hilbert H.,  
 RA Borzom K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Medler H., Wambitz R., Punnett B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaur V., Mortier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Roche M., Galliard C., Talada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forburg S.L.,  
 RA Cirutici L., Lowe T., McCombie W.R., Paulsen I., Porschke J.,  
 RA Sipakovski G.V., Ussery D., Barrett B.G., Nurse P.,  
 RT "the genome sequence of Schizosaccharomyces pombe";  
 RT Nature 415:871-880(2002).  
 CC -1- CATALYTIC ACTIVITY: Acyl-CoA + cholesterol = CoA + cholesterol  
 CC ester.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic  
 CC reticulum (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE STEROL O-ACYLTRANSFERASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: Z69729; CA93593.1; -.  
 DR PIR: S67434; S67434.  
 DR GeneDB: Spombe; SPAC1367.05; -.  
 DR Pfam: PF03062; MBOAT; 1.  
 DR KMW: Hypothetical protein; Transferrase; Transmembrane;  
 KW Endoplasmic reticulum; Acyltransferase.  
 FT TRANSMEM 140 118 POTENTIAL.  
 FT TRANSMEM 174 194 POTENTIAL.  
 FT TRANSMEM 199 219 POTENTIAL.  
 FT TRANSMEM 305 325 POTENTIAL.  
 FT TRANSMEM 344 364 POTENTIAL.  
 FT TRANSMEM 384 404 POTENTIAL.  
 FT TRANSMEM 462 482 POTENTIAL.  
 FT TRANSMEM 488 508 POTENTIAL.  
 FT TRANSMEM 517 537 POTENTIAL.  
 FT CARBOHYD 250 250 POTENTIAL.  
 SQ SEQUENCE 537 AA; 63098 MW; 41DBA79402B67A28 CRC64; .  
 QY Query Match 62.3%; Score 38; DB 1; Length 537;  
 Best Local Similarity 58.3%; Pred. No. 18;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 PLSISFRIQDP 12  
 DB 79 PRKSIFDRVTPD 90  
 RESULT 12  
 ID VVRA CHLTR STANDARD; PRT; 1786 AA.  
 AC 084337;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE UVRAIC system protein A (Uvra protein) (Excinuclease ABC subunit A).  
 GN UVRA OR CT333.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UW-3/Cx;  
 RX MEDLINE=99000809; PubMed=9784136.  
 RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
 RA Mitchell W.P., Olinger L., Tatunov R.L., Zhao Q., Koonin E.V.,  
 RA Davis R.W.;

RT "Genome sequence of an obligate intracellular pathogen of humans:  
 RT Chlamydia trachomatis.";  
 RT Science 282:754-759(1998).  
 CC -1- FUNCTION: The UvrABC repair system catalyzes the recognition and  
 CC processing of DNA lesions. UvrA is an ATPase and a DNA-binding  
 CC protein. A damage recognition complex composed of 2 uvrA and 2  
 CC uvrB subunits scans DNA for abnormalities. When the presence of a  
 CC lesion has been verified by uvrB, the uvrA molecules dissociate  
 CC (By similarity).  
 CC -1- SUBUNIT: Forms a heterotrimer with uvrB during the search for  
 CC lesions (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. Uvra SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE001306; AAC67928.1; -.  
 DR PIR: H71527; H71527.  
 DR HAMAP: MF\_00205; -; 2.  
 DR InterPro: IPR003439; ABC transporter.  
 DR InterPro: IPR004602; UvrA.  
 DR Pfam: PF00005; ABC\_tran; 2.  
 DR ProDom: PD000006; ABC\_transporter; 1.  
 DR TIGRfam: TIGR00630; UvrA; 1.  
 DR DR PROSITE: PS00211; ABC\_TRANSPORTER\_1; FALSE\_NEG.  
 DR PROSITE: PS50893; ABC\_TRANSPORTER\_2; 2.  
 KW SOS response; Excision nuclease; DNA repair; DNA recombination;  
 KW DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;  
 KW Zinc-finger; Complete proteome.  
 FT NP\_BIND 32 39 ATP (POTENTIAL).  
 FT NP\_BIND 625 632 ATP (POTENTIAL).  
 FT ZN\_FING 719 742 C4-TYPE.  
 FT NP\_BIND 964 971 ATP (POTENTIAL).  
 FT NP\_BIND 1516 1523 ATP (POTENTIAL).  
 FT ZN\_FING 1602 1628 C4-TYPE.  
 SQ SEQUENCE 1786 AA; 196948 MW; 02D6862EE15DE070 CRC64;  
 QY Query Match 62.3%; Score 38; DB 1; Length 1786;  
 Best Local Similarity 63.6%; Pred. No. 63;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 LSSIFSRIGDP 12  
 DB 1042 LESLFTRIQHP 1052  
 RESULT 13  
 ID YOR2 CAEEL STANDARD; PRT; 495 AA.  
 AC 009307;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Hypothetical 45.1 kDa protein f19c6.2 in chromosome X.  
 GN f19c6.2.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Harris B.;  
 RA Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; 248006; CNA8050.1; -.  
DR PIR; T2115; T2115.  
DR WormRep; P19C6.2; CE032231.  
KW Hypothetical protein.  
SQ SEQUENCE 495 AA; 55814 MW; 99EB952032B0B48C CRC64;  
  
Query Match 60.7%; Score 37; DB 1; Length 495;  
Best Local Similarity 45.5%; Pred. No. 26;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 PUSISFRRIGD 11  
Db 278 PVSGLFQKVD 288  
  
RESULT 14  
MWIN SYN3 STANDARD; PRT; 533 AA.  
ID MWIN SYN3  
AC 055179;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Virulence factor mvin homolog.  
GN MWIN OR SLR0488.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_Taxid=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96127529; PubMed=8590279;  
RA Kaneo T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,  
RA Sugita M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb  
RT region from map positions 644 to 924 of the genome.";  
RL DNA Res. 2:153-166(1995).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: BELONGS TO THE MWIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D64001; BAA10319.1; -.  
DR PIR; S74401.  
DR InterPro; IPR004268; MWIN\_1like.  
DR Pfam; PF03023; MWIN; 1.  
KW Transmembrane; Complete proteome.  
SQ SEQUENCE 533 AA; 57053 MW; 00ASB612F744CID2 CRC64;

Query Match 60.7%; Score 37; DB 1; Length 533;  
Best Local Similarity 50.0%; Pred. No. 28;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
  
Qy 1 PUSISFRRIGD 12  
Db 300 PLPVFSLRADP 311  
  
RESULT 15  
S2IG\_HUMAN STANDARD; PRT; 848 AA.  
ID S2IG\_HUMAN  
AC 09H2Y9;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Solute carrier family 21 member 15 (Organic anion transporter  
DE polypeptide-related protein 4) (OATP-RP4) (OATRP4).  
GN SLC21A15.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Wu Y., Heiang B.H., Zhu Y., Yang W.-P., Kirchgessner T.G.;  
RT "Identification and characterization of novel human OATP family  
RT members.";  
RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RL SLC21A15 LOCATOR: Integral membrane protein.  
CC -1- SIMILARITY: BELONGS TO THE SLC21 FAMILY OF TRANSPORTERS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF205075; AAG42207.1; -.  
DR Genew; HGNC:19046; SLC21A15.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR004157; OATP\_Cterm.  
DR InterPro; IPR004156; OATP\_Nterm.  
DR Pfam; PF03137; OATP\_C; 1.  
DR Pfam; PF03132; OATP\_N; 1.  
DR TIGRPFAM; TIGR00805; oat; 1.  
KW Transmembrane; Transport; Glycoprotein.  
SQ SEQUENCE 129 AA; 149 MW; 188 MW; 188 MW;  
FT TRANSMEM 168 188 POTENTIAL.  
FT TRANSMEM 198 218 POTENTIAL.  
FT TRANSMEM 261 281 POTENTIAL.  
FT TRANSMEM 304 324 POTENTIAL.  
FT TRANSMEM 348 368 POTENTIAL.  
FT TRANSMEM 435 455 POTENTIAL.  
FT TRANSMEM 473 493 POTENTIAL.  
FT TRANSMEM 506 526 POTENTIAL.  
FT TRANSMEM 651 671 POTENTIAL.  
FT TRANSMEM 696 716 POTENTIAL.  
FT TRANSMEM 736 756 POTENTIAL.  
FT DOMAIN 377 383 POLY-LYS.  
FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 431 431 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 550 550 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 590 590 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 597 597 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 730 730 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQL SEQUENCE 848 AA; 91877 MW; F7BDC260744B7FA4 CRC64;

Query Match 60.7%; Score 37; DB 1; Length 848;

Best Local Similarity 50.0%; Pred. NO. 45;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LSSISFRIGDP 12

Db 776 PLSTVSERVGHF 787

RESULT 16

DCUP\_CHLMU STANDARD; PRT; 334 AA.

AC 09PLH7;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Uroporphyrinogen decarboxylase (EC 4.1.1.37) (URO-D) (UPD).

DE HEME OR TC0123.

OS Chlamydia muridarum.

OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI\_Taxid=83560;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MOPN / N199;

RX MEDLINE=20150255; PubMed=10684935;

RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,

White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,

Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,

RA Gaim M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,

Eisen J., Fraser C.M.,

RT "Genome sequences of Chlamydia trachomatis Mohn and Chlamydia

pneumoniae AR3."

RL Nucleic Acids Res. 28:1397-1406(2000).

CC -1- CATALYTIC ACTIVITY: Uroporphyrinogen-III = coproporphyrinogen + 4

CO(2).

CC -1- PATHWAY: Porphyrin biosynthesis.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -1- SIMILARITY: BELONGS TO THE UROPORPHYRINOGEN DECARBOXYLASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AE002279; AAF39001.1; -

DR PIR; F81739; F81739.

DR HSSP; P06132; URO.

DR TIGR; TC0123; -

DR HAMAP; MF\_00218; -; 1.

DR InterPro; IPR006361; Heme.

DR InterPro; IPR000257; Uro-decarboxyls.

DR Pfam; PF01208; URO-D; 1.

DR ProDom; PD003225; Uro-decarboxyls; 1.

DR TIGRfam; TIGR01464; Heme; 1.

DR PROSITE; PS00906; UROD\_1; 1.

DR PROSITE; PS00907; UROD\_2; 1.

KW Lyase; Decarboxylase; Porphyrin biosynthesis; Complete proteome.

SM SEQUENCE 334 AA; 37380 MW; 199BA88AC1F4DD CRC64;

QY 2 LSSISFRIGDP 12

Db 258 LSEIYRQIGDP 268

RESULT 17

ID VMSA\_HPBW

AC P03142; STANDARD; PRT; 369 AA.

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Major surface antigen precursor.

GN S.

OS Hepatitis B virus (subtype adw2).

OC Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.

OX NCBI\_Taxid=106821;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=83168919; PubMed=6300776;

RA Ono Y., Ono H., Sasada R., Igarashi K., Sugino Y., Nishio K.,

RT "The complete nucleotide sequences of the cloned hepatitis B virus

RT DNA; subtype adr and adw."

RL Nucleic Acids Res. 11:1747-1757(1983).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; V00866; CA24233.1; ALT\_INIT.

DR PIR; A93460; SALVE.

DR InterPro; IPR000349; Hepvir\_surfag.

DR Pfam; PF00695; VMSA; 1.

KM Antigen.

FT PROPP 1 163

FT CHAIN 164 389

FT CARBOHYD 4 4

FT CARBOHYD 26 26

FT CARBOHYD 112 112

FT CARBOHYD 166 166

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

FT CARBOHYD 309 309

SQL SEQUENCE 389 AA; 42354 MW; FB9EF27BEC5B5 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 389;

Best Local Similarity 63.6%; Pred. NO. 31;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LSSISFRIGDP 12

Db 150 ISSISARTGDP 160

RESULT 18

ID VMSA\_HPBW2

AC P03141; STANDARD; PRT; 400 AA.

DT 21-JUL-1986 (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Major surface antigen precursor.

GN S.

OS Hepatitis B virus (subtype adw2).

OC Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.

OX NCBI\_Taxid=10408;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=79244739; PubMed=471053;

RA Valenzuela P., Jaenisch R., Fox C.F. (eds.);

RT Animal virus genetics, pp.57-70, Academic Press, New York (1980).

RL [2]

RP SEQUENCE OF 175-400 FROM N.A.

RX MEDLINE=79244739; PubMed=471053;

RA Valenzuela P., Gray P., Quiroga M., Zaldívar J., Goodman H.M.,

RT Rutter W.J.;



RT "Nucleotide sequence of the gene coding for the major protein of  
hepatitis B virus surface antigen.";  
RL Nature 280:815-819 (1979).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X02763; CAA26539.1; -  
DR EMBL; J02205; AAA4524.1; -  
DR PIR; A03706; SAVLYD.  
DR InterPro; IPR000349; Hepvir\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
FT PROPEP 1 174  
FT CHAIN 175 400 MAJOR SURFACE ANTIGEN.  
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 400 AA; 43704 MW; 57356B6293872BC5 CRC64;  
Query Match 59.0%; Score 36; DB 1; Length 400;  
Best Local Similarity 63.6%; Pred. No. 32;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 LSISFRIGDP 12  
DB 161 ISSISARTGDP 171  
RESULT 19  
VMSA\_HPBV9 STANDARD; PRT; 400 AA.  
ID VMSA\_HPBV9 17101;  
AC P17101;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 01-AUG-1992 (Rel. 23, Last annotation update)  
DE Major surface antigen precursor.  
GN S.  
OS Hepatitis B virus (subtype adw / strain 991).  
OC Viruses; Retrovird viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10410;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Koechel H.G., Schueler A., Lotmann S., Thomassen R.;  
RL Submitted (FEB-1990) to the EMBL/GenBank/DBD databases.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X51970; CAA36230.1; -  
DR PIR; S10383; SAVLKS.  
DR InterPro; IPR000349; Hepvir\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
FT PROPEP 1 174  
FT CHAIN 175 400 MAJOR SURFACE ANTIGEN.  
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 400 AA; 43695 MW; 76D98F74B8B17724 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 400;  
Best Local Similarity 63.6%; Pred. No. 32;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 LSISFRIGDP 12  
DB 161 ISSISARTGDP 171  
RESULT 20  
MP44\_MCV1 STANDARD; PRT; 593 AA.  
ID MP44\_MCV1 098224; O11329;  
AC 098224; O11329;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Probable metalloendopeptidase Gl-type (EC 3.4.24.-).  
GN MC056L OR B2'-17R.  
OS Molluscum contagiosum virus subtype 1 (MCV1).  
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;  
OX Molluscipoxvirus.  
CC NCBI\_TaxID=10280;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96325459; PubMed=8670425;  
RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darel G.,  
RT Mose B.;  
RT "Genome sequence of a human tumorigenic poxvirus: prediction of  
RT specific host response-evasion genes.";  
RL Science 273:813-816 (1996).  
RN [2]  
RP SEQUENCE OF 206-392 FROM N.A.  
RA Moratilla M., Agromayor M., Nunez A., Funes J.M., Varas A.J.,  
RA Lopez-Esteban J.L., Esteban M., Martin-Gallardo A.;  
RT "A random DNA sequencing, computer-based approach for the  
RT generation of a gene map of Molluscum Contagiosum virus.";  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBD databases.  
CC  
CC -1- FUNCTION: Seems to be involved in viral proteins maturation by  
CC cleavage at Ala-Gly--Xaa motifs (by similarity).  
CC  
CC -1- COTACTOR: BINDS 1 ZINC ION (Potential).  
CC  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M44.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U60315; AAC55184.1; -  
DR EMBL; U66909; AAB57961.1; -  
DR PIR; T30658; T30658.  
DR MEROPS; M44.001; -  
DR InterPro; IPR005072; Peptidase\_M44.  
DR Pfam; PF03410; Peptidase\_M44; 1.  
KW Hydrolase; Metalloprotease; Zinc.  
FT METAL 41 41 ZINC (CATALYTIC) (POTENTIAL).  
FT ACT SITE 44 44 ZINC (CATALYTIC) (POTENTIAL).  
FT METAL 45 45 ZINC (CATALYTIC) (POTENTIAL).  
SQ SEQUENCE 593 AA; 68419 MW; 142C620E1AB0C61B CRC64;  
Query Match 59.0%; Score 36; DB 1; Length 593;  
Best Local Similarity 50.0%; Pred. No. 48;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
QY 1 PLSIFSRIGDP 12  
DB 381 PRAEVRVRIKDP 392  
RESULT 21  
EFGI\_SVNY3

ID EFG1, SYNY3 STANDARD; PRT; 695 AA.  
AC P28371; P74556;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Elongation factor G 1 (EF-G 1).  
OS FUSA OR FUS OR SLR1463.  
GN Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
NCBI\_TaxID=1148;  
RX MEDLINE=94339485; PubMed=8061323;  
RA Melisch P.L., Johnson D., Zhang Y., Breitenberger C.A.;  
RA "Synechocystis" sp. PCC6803 fusi gene, located outside of the str  
RT operon, encodes a polypeptide related to protein synthesis factor  
RT EF-G.";  
RL Plant Mol. Biol. 25:735-738 (1994).  
[2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kocani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirosewa M., Sugita M., Sasanoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Natsu K.,  
RA Okumura S., Shimo S., Takeuchi C., Wada T., Watanabe A.,  
RA Yamada M., Yasuda M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136 (1996).  
CC - FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLLOCATION  
CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE  
CC RIBOSOME.  
CC - SUBCELLULAR LOCATION: Cytoplasmic.  
CC - SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
CC EF-G/EF-2 SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X65159; CAA46277.1; -;  
CC EMBL; D90916; BAI18663.1; -;  
CC PIR; S76751; S76751.  
CC DR HSP; P13551; 2BFG.  
CC DR HAMAP; MF\_00054; -; 1.  
CC DR InterPro; IPR004540; EF-G.  
CC DR InterPro; IPR000795; EF\_GTPbind.  
CC DR InterPro; IPR000640; EFG\_C.  
CC DR InterPro; IPR005517; EFG\_IV.  
CC DR InterPro; IPR004161; EFTU\_D2.  
CC DR InterPro; IPR005225; Small\_GTP.  
CC DR Pfam; PF00679; EFG\_C; 1.  
CC DR Pfam; PF03764; EFG\_IV; 1.  
CC DR Pfam; PF00009; GTP\_EFTU; 1.  
CC DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
CC DR PRINTS; PR00315; ELONGATNFACT.  
CC DR TIGRFAMs; TIGR00484; EF-G; 1.  
CC DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
CC DR PROSITE; PS00301; EFACOR\_GTP; 1.  
CC DR Elongation factor; protein biosynthesis; GTP-binding;  
CC MultiGene family; Complete proteome.  
CC KW NE\_BIND 15  
CC FT NE\_BIND 79 GTP (BY SIMILARITY).  
CC FT NE\_BIND 83 GTP (BY SIMILARITY).  
CC FT NE\_BIND 133 GTP (BY SIMILARITY).  
CC FT CONFICT 42 H->R (IN REF. 1).  
CC SQ SEQUENCE 695 AA; 76750 MW; 8727830BP9073223 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 695;

Best Local Similarity 63.6%; Pred. No. 56;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 PLUSIFSRIGD 11  
DB 645 PLSMFYIGD 655  
RESULT 22  
EFG\_MICLU STANDARD; PRT; 701 AA.  
AC P09352;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Elongation factor G (EF-G).  
GN FUSA OR FUS.  
OS Micrococcus luteus (Micrococcus lysodeikticus).  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Micrococccinae; Micrococccaceae; Micrococcus.  
NCBI\_TaxID=1270;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88007427; PubMed=3654584;  
RA Ohama T., Yamao F., Muto A., Osawa S.;  
RT "Organization and codon usage of the streptomycin operon in  
RT Micrococcus luteus, a bacterium with a high genomic G + C content.";  
RL J. Bacteriol. 169:4770-4777 (1987).  
CC - FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLLOCATION  
CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE  
CC RIBOSOME.  
CC - SUBCELLULAR LOCATION: Cytoplasmic.  
CC - SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
CC EF-G/EF-2 SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M17788; AAA25319.1; -;  
CC PIR; C26956; C26956.  
CC DR HSP; P13551; 1DAR.  
CC DR HAMAP; MF\_00054; -; 1.  
CC DR InterPro; IPR004540; EF-G.  
CC DR InterPro; IPR000795; EF\_GTPbind.  
CC DR InterPro; IPR000640; EFG\_C.  
CC DR InterPro; IPR005517; EFG\_IV.  
CC DR InterPro; IPR004161; EFTU\_D2.  
CC DR InterPro; IPR005225; Small\_GTP.  
CC DR Pfam; PF00679; EFG\_C; 1.  
CC DR Pfam; PF03764; EFG\_IV; 1.  
CC DR Pfam; PF00009; GTP\_EFTU; 1.  
CC DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
CC DR PRINTS; PR00315; ELONGATNFACT.  
CC DR TIGRFAMs; TIGR00484; EF-G; 1.  
CC DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
CC DR PROSITE; PS00301; EFACOR\_GTP; 1.  
CC DR Elongation factor; protein biosynthesis; GTP-binding;  
CC KW NE\_BIND 15  
CC FT NE\_BIND 79 GTP (BY SIMILARITY).  
CC FT NE\_BIND 83 GTP (BY SIMILARITY).  
CC FT NE\_BIND 133 GTP (BY SIMILARITY).  
CC SQ SEQUENCE 701 AA; 77425 MW; 25E49FACF350E236 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 701;  
Best Local Similarity 63.6%; Pred. No. 57;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 PLUSIFSRIGD 11

Db 655 PLSMFYIGD 665

RESULT 23

MUTS\_VIBCH STANDARD; PRT; 862 AA.

AC 09KUI6;  
 DT 15-SEP-2003 (Rel. 42, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE DNA mismatch repair protein muts.  
 GN MUTS OR VC0535.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrrio.  
 OX NCBI\_TaxId=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=El Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae.";  
 RL Nature 406:477-483(2000).  
 CC -1- FUNCTION: This protein is involved in the repair of mismatches in  
 CC DNA. It is possible that it carries out the mismatch recognition  
 CC step. This protein has a weak ATPase activity (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE004139; AAF93703.1; -  
 CC TIGR; VC0535; -  
 DR HAMAP; MF\_00096; -; 1.  
 DR InterPro; IPR005748; MutS1.  
 DR InterPro; IPR000432; MutS\_C.  
 DR InterPro; IPR007860; MutS\_II.  
 DR InterPro; IPR007696; MutS\_III.  
 DR InterPro; IPR007861; MutS\_IV.  
 DR InterPro; IPR007695; MutS\_N.  
 DR Pfam; PF01624; MutS\_1; 1.  
 DR Pfam; PF05188; MutS\_II; 1.  
 DR Pfam; PF05192; MutS\_III; 1.  
 DR Pfam; PF05190; MutS\_IV; 1.  
 DR Pfam; PR00488; MutS\_V; 1.  
 DR ProDom; PD001263; MutS\_C; 1.  
 DR SMART; SM00534; MUTSAC; 1.  
 DR SMART; SM00533; MUTSD; 1.  
 DR TIGRfam; TIGR01070; muts1; 1.  
 DR PROSITE; PS00486; DNA\_MISMATCH\_REPAIR\_2; 1.  
 DR DNA repair; ATP-binding; DNA-binding; Complete proteome.  
 KW NP\_BIND 621 628 ATP (POTENTIAL).  
 FT SEQUENCE 862 AA; 96326 MW; 98882509B7140531 CRC64;  
 SQ

Query Match 59.0%; Score 36; DB 1; Length 862;  
 Best Local Similarity 70.0%; Pred. No. 70;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PLSIFSRIG 10  
 DB 656 PLDIRIFRIG 665

RESULT 24

PDR6\_YEAST STANDARD; PRT; 1081 AA.

AC P32767;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Pleiotropic drug resistance regulatory protein 6.  
 GN PDR6 OR KAP122 OR YGL016W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OS Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxId=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Il125-2B;  
 RX MEDLINE=91353083; PubMed=1882553;  
 RA Chen W., Balzi E., Capleau E., Choder M., Goffeau A.;  
 RT "The DNA sequencing of the 17 kb HindIII fragment spanning the LEU1  
 RT and ATR1 loci on chromosome VII from Saccharomyces cerevisiae reveals  
 RT pleiotropic drug resistance.";  
 RL Yeast 7:287-299(1991).  
 CC [2]  
 RP SEQUENCE FROM N.A.  
 RA Hedling U., Hofmann B., Delius H.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: PROBABLE POSITIVE REGULATOR OF PROTEINS INVOLVED IN  
 CC PERMEABILITY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; S58126; AADI3901.1; -  
 CC EMBL; S57895; AABI9613.1; -  
 DR EMBL; Z72538; CAA96716.1; -  
 DR PIR; S15040; S15040.  
 DR SGD; S0002984; KAP122.  
 DR GO; GO:0005737; C:cytoplasm; IDA.  
 DR GO; GO:0005643; C:nuclear pore; IDA.  
 DR GO; GO:0006606; P:protein-nucleus import; IDA.  
 SQ SEQUENCE 1081 AA; 123530 MW; EBBCSADD4E16D9D3 CRC64;  
 SQ

Query Match 59.0%; Score 36; DB 1; Length 1081;  
 Best Local Similarity 63.6%; Pred. No. 89;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LSLIFKINDP 12  
 DB 118 LSLIFKINDP 128

RESULT 25

EFG\_STRRA STANDARD; PRT; 341 AA.

AC P29541;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Elongation factor G (EF-G) (Fragment).  
 GN FUS.  
 OS Streptomyces ramocissimus.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Streptomycinae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxId=1925;  
 RN [1]

```

RP SEQUENCE FROM N.A.
RX MEDLINE=94282341; PubMed=8012612;
RA van Vlijmenboom E., Moudt L.P., Heinstra P.W.H., Rietveld K.,
RT "Three tuf-like genes in the kitomyacin producer Streptomyces
RT ramocissimus.";
RL Microbiology 140:983-998(1994).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLLOCATION
CC OR THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
CC RIBOSOME.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X67057; CAA47441.1; -.
DR PIR; S23907; S23907.
DR HSSP; P13551; IELQ.
DR HAMAP; MF_00054; -.
DR InterPro; IPR000795; PF_GTPbind.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR005517; EFG_IV.
DR InterPro; IPR004161; EFTU_D2.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF03764; EFG_IV; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PROSITE; PS00301; EFATOR_GTP; PARTIAL.
DR Elongation factor; Protein biosynthesis; GTP-binding.
KW NON TER
FT 1
SQ SEQUENCE 341 AA; 37322 MW; DEEDCC72A0A62F91 CRC64;

Query Match 57.4%; Score 35; DB 1; Length 341;
Best Local Similarity 54.5%; Pred. No. 42;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PLSISFRIGDP 11
DB 295 PLSMFVGVGD 305

RESULT 26
VMSA_HPBVI STANDARD; PRT; 389 AA.
AC P17397;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface antigen precursor.
GN S
OS Hepatitis B virus (subtype adw / strain Indonesia/PIIM420).
OS Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10412;
RX MEDLINE=89010694; PubMed=3171552;
RA Okamoto H., Tsuda F., Sakugawa H., Saetosewignjo R.I., Imai M.,
RA Miyakawa Y., Mayumi M.;
RT "Typing hepatitis B virus by homology in nucleotide sequence:
RT comparison of surface antigen subtypes.";
RL J. Gen. Virol. 69:2575-2583(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D00331; -. NOT_ANNOTATED_CDS.
DR PIR; I28925; SAVLJ3.
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT PROPEP 1 163
FT CHAIN 164 389
FT CARBOHYD 4 4
FT CARBOHYD 112 112
FT CARBOHYD 166 166
FT CARBOHYD 309 309
SQ SEQUENCE 389 AA; 42653 MW; 08079BA34F3B9C0 CRC64;

Query Match 57.4%; Score 35; DB 1; Length 389;
Best Local Similarity 63.6%; Pred. No. 48;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LSSIFSRIGDP 12
DB 150 ISSILSTGDP 160

RESULT 27
VMSA_HPBVT STANDARD; PRT; 400 AA.
AC Q05496;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Major surface antigen precursor.
GN S.
OS Hepatitis B virus (subtype adw4 / strain Brazil / isolate w4B).
OS Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=5410;
RX MEDLINE=93346970; PubMed=8345355;
RA Naumann H., Schaefer S., Yoshida C.F.T., Gaspar A.M.C., Repp R.,
RA Gerlich W.H.;
RT "Identification of a new hepatitis B virus (HBV) genotype from Brazil
RT that expresses HBV surface antigen subtype adw4.";
RL J. Gen. Virol. 74:1627-1632(1993).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X69798; CAA49455.1; -.
DR PIR; J02230.
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT PROPEP 1 174
FT CHAIN 175 400
SQ SEQUENCE 400 AA; 43551 MW; 4A5A2212E4B3E117 CRC64;

Query Match 57.4%; Score 35; DB 1; Length 400;
Best Local Similarity 70.0%; Pred. No. 49;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 SSIFSRIGDP 12
DB 162 SSIFSKTGP 171

RESULT 28

```

```

IPGD_SHIFL STANDARD; PRT; 538 AA.
ID IPGD_SHIFL
AC Q07566; Q8VSH4;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE IPGD protein.
GN IPGD OR CP0133.
OS Shigella flexneri.
OC Plasmid pWR100, and Plasmid pCP301.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M90T / Serotype 5; PLASMID=pWR100;
RX MEDLINE=93339268; PubMed8478058;
RA Allouat A., Menard R., Sansonetti P.J., Parrot C.;
RT "Characterization of the Shigella flexneri ipgd and ipgf genes, which
are located in the proximal part of the mxi locus.";
RL Infect. Immun. 61:1707-1714(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=M90T / Serotype 5; PLASMID=pWR100;
RX MEDLINE=20566792; PubMed1115111;
RA Buchrieser C., Glaser P., Rusnok C., Nedjari H., d'Hauteville H.,
Kunat F., Sansonetti P., Parrot C.;
RT "The virulence plasmid pWR100 and the repertoire of proteins secreted
by the type III secretion apparatus of Shigella flexneri.";
RL Mol. Microbiol. 38:760-771(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=M90T / Serotype 5; PLASMID=pWR100;
RX MEDLINE=21189246; PubMed11292750;
RA Venkatesan M.M., Goldberg M.B., Rose D.J., Grotbeck E.J., Burland V.,
Blattner F.R.;
RT "Complete DNA sequence and analysis of the large virulence plasmid of
Shigella flexneri.";
RL Infect. Immun. 69:3271-3285(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a; PLASMID=pCP301;
RX MEDLINE=22272406; PubMed12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
Sun J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L04309; AAA26517.1; -
DR EMBL; AL391753; CAC05808.1; -
DR EMBL; AF348706; AAK18452.1; -
DR EMBL; AF36526; AAL72339.1; -
KM Virulence; Plasmid.
FT VARIANT 73 73 K -> R (IN PLASMID PCP301).
FT VARIANT 155 155 N -> D (IN PLASMID PCP301).
FT VARIANT 166 166 S -> G (IN PLASMID PCP301).
FT VARIANT 200 200 R -> S (IN PLASMID PCP301).
FT VARIANT 300 300 S -> N (IN PLASMID PCP301).
FT VARIANT 311 311 G -> E (IN PLASMID PCP301).
FT VARIANT 406 406 NN -> TK (IN PLASMID PCP301).
FT VARIANT 424 424 V -> M (IN PLASMID PCP301).

```

```

SQ SEQUENCE 538 AA; 59832 MW; 304D27839C64CF96 CRC64;
Query Match
Best Local Similarity 56.2%; Score 35; DB 1; Length 538;
Matches 9; Conservative 1; Mismatches 2; Indels 4; Gaps 1;
Qy 1 PLSST-----FSRIQDP 12
|||: |||
Db 510 PLSSTLSYSERIGDP 525

RESULT 29
EFG_MYCLE STANDARD; PRT; 701 AA.
ID EFG_MYCLE
AC P30767;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation factor G (EF-G).
GN FUS4 OR EFG OR ML1878.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93188701; PubMed8446028;
RA Honore N.T., Berth S., Chanteau S., Doucet-Populaire F.,
Rigamonti K., Garnier T., Georges C., Lannots P., Limpaiboon T.,
Newton S., Niang K., del Portillo P., Ramesh G.R., Reddi P.,
Ridel P.R., Sittisombut N., Wu-Hunter S., Cole S.T.;
RT "Nucleotide sequence of the first cosmid from the Mycobacterium
leprae genome project: structure and function of the Rlf-Str
regions.";
RL Mol. Microbiol. 7:207-214(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed11234002;
RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
Davies R.M., Deavin K., Duthoy S., Feltingwell T., Fraser A., Hamlin N.,
Holroyd S., Hornby T., Jagsels K., Lacroix C., Maclean J., Moule S.,
Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
Rutter S., Seeger K., Simon S., Simmonds M., Skellon J., Squares R.,
Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -----
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
CC RIBOSOME.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 214314; CAA78673.1; -
DR EMBL; AL583923; CAC30832.1; -
DR PIR; H87143; H87143.
DR PIR; S31150; S31150.
DR HSRP; P13511; IEFNM.
DR Leproma; ML1878; -.
DR HAMAP; MF_00054; -; 1.
DR InterPro; IPR004540; EF-G.

```

```
DR InterPro: IPR000795; EF_GTPbind.
DR InterPro: IPR000640; EFG_C.
DR InterPro: IPR005517; EFG_IV.
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00679; EFG_C; 1.
DR Pfam: PF03764; EFG_IV; 1.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR PRINTS: PRO0315; ELONGATINFC.
DR TIGRPFAM: TIGR00484; EF-G; 1.
DR TIGRPFAM: TIGR00231; small_GTP; 1.
DR PROSITE: PS00301; EFACITOR_GTP; 1.
DR Elongation factor: Protein biosynthesis; GTP-binding;
KM Complete proteome.
FT NP_BIND 20 27 GTP (BY SIMILARITY).
FT NP_BIND 84 88 GTP (BY SIMILARITY).
FT NP_BIND 138 141 GTP (BY SIMILARITY).
FT CONFLICT 564 564 G -> A (IN REF. 1).
SQ SEQUENCE 701 AA; 77160 MW; F18B724F35F3181E CRC64;

Query Match 57.4%; Score 35; DB 1; Length 701;
Best Local Similarity 54.5%; Pred. No. 88;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 PLSSIFSRIGD 11
DB 655 PUSEMFGYVD 665

RESULT 30
EFG_MYCTU STANDARD; PRT; 701 AA.
AC 053790;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation factor G (EF-G).
GN FUSA OR RV0684 OR MT0712 OR MT040.12 OR MTCV210.01.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Mycobacterium.
OX NCBI_Taxid=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Broesch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultsen J.E., Taylor K., Whitehead S., Barrett B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Petersen J., DeBoy R., Dodson R., Gwin M.L., Hatt D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Mayhew L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLATION
CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
CC RIBOSOME.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
```

```
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL021943; CAI17471.1; -.
CC EMBL: AE006965; AAK44938.1; -.
CC PIR: E70827; E70827.
CC HSSP: P13551; IFNM.
CC TIGR: MT0712; -.
CC Tuberculosis; RV0684; -.
CC HAMAP: MF_00054; -.
CC InterPro: IPR004540; EF-G.
CC InterPro: IPR000795; EF_GTPbind.
CC InterPro: IPR005517; EFG_C.
CC InterPro: IPR004161; EFTU_D2.
CC InterPro: IPR005225; Small_GTP.
CC Pfam: PF00679; EFG_C; 1.
CC Pfam: PF03764; EFG_IV; 1.
CC Pfam: PF00009; GTP_EFTU; 1.
CC Pfam: PF03144; GTP_EFTU_D2; 1.
CC PRINTS: PRO0315; ELONGATINFC.
CC TIGRPFAM: TIGR00484; EF-G; 1.
CC TIGRPFAM: TIGR00231; small_GTP; 1.
CC PROSITE: PS00301; EFACITOR_GTP; 1.
CC Elongation factor: Protein biosynthesis; GTP-binding;
KM Complete proteome.
FT NP_BIND 20 27 GTP (BY SIMILARITY).
FT NP_BIND 84 88 GTP (BY SIMILARITY).
FT NP_BIND 138 141 GTP (BY SIMILARITY).
SQ SEQUENCE 701 AA; 77202 MW; 0C6DD4C128E31EAL CRC64;

Query Match 57.4%; Score 35; DB 1; Length 701;
Best Local Similarity 54.5%; Pred. No. 88;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 PLSSIFSRIGD 11
DB 655 PUSEMFGYVD 665

RESULT 31
EFG_CORGL STANDARD; PRT; 705 AA.
AC 08NT19;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation factor G (EF-G).
GN FUSA OR CGJ0495.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinomycetales;
OC Corynebacteriaceae; Corynebacterium.
OX NCBI_Taxid=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: This protein promotes the GTP-dependent translocation of
CC the nascent protein chain from the A-site to the P-site of the
CC ribosome.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.
```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AP005275; BAB97888.1; -.
DR HAMAP, MF_00054; -.
DR InterPro; IPR004540; EF-G.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR005517; EFG_IV.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF03764; EFG_IV; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PRO0315; ELONGATNFCT.
DR TIGRfams; TIGR00484; EF-G; 1.
DR TIGRfams; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFACITOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding;
KW Complete proteome.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 79 83 GTP (BY SIMILARITY).
FT NP_BIND 133 136 GTP (BY SIMILARITY).
SQ SEQUENCE 705 AA; 77464 MW; 6CD7010C4110B2F2 CRC64;

Query Match 57.4%; Score 35; DB 1; Length 705;
Best Local Similarity 54.5%; Pred. No. 88;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 PLSSTFSRIGD 11
DB 657 PLSQMFQVGD 667

RESULT 32
ID EFG1_STRCO STANDARD; PRT; 708 AA.
AC P40173; OSL0K3;
DT 01-FEB-1995 (Rel. 31, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Elongation factor G 1 (EF-G 1).
GN FUS1 OR FUS OR SCO4661 OR SCO8404.07.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycinae; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=A3(2) / M145;
RC MEDLINE=2196410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kiese H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kiese T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)";
RL Nature 417:141-147(2002).
RN [2]
RP SEQUENCE OF 616-708 FROM N.A.
RC STRAIN=A3(2) / M145;

```

```

RX MEDLINE=95002174; PubMed=7918656;
RA van Wezel G.P., Moud L.P., Verweene R., Verdurmen M.L.,
RA Vijgenboom E., Bosch L.;
RT "Cloning and sequencing of the tuf genes of Streptomyces coelicolor
RT A3(2)";
RL Biochim. Biophys. Acta 1219:543-547(1994).
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLOCATION
CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE
CC RIBOSOME.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC EF-G/EF-2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AL939121; CAB81852.1; -.
DR EMBL; X77039; CA54328.1; -.
DR PIR; S50137; S50137.
DR HSSP; P13551; 1DAR.
DR HAMAP; MF_00054; -.
DR InterPro; IPR004540; EF-G.
DR InterPro; IPR000795; EF_GTPbind.
DR InterPro; IPR000640; EFG_C.
DR InterPro; IPR005517; EFG_IV.
DR InterPro; IPR004161; EFTU_D2.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00679; EFG_C; 1.
DR Pfam; PF03764; EFG_IV; 1.
DR Pfam; PF00009; GTP_EFTU; 1.
DR Pfam; PF03144; GTP_EFTU_D2; 1.
DR PRINTS; PRO0315; ELONGATNFCT.
DR TIGRfams; TIGR00484; EF-G; 1.
DR TIGRfams; TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFACITOR_GTP; 1.
KW Elongation factor; Protein biosynthesis; GTP-binding;
KW Multigene family; Complete proteome.
FT NP_BIND 18 25 GTP (BY SIMILARITY).
FT NP_BIND 86 90 GTP (BY SIMILARITY).
FT NP_BIND 140 143 GTP (BY SIMILARITY).
SQ SEQUENCE 708 AA; 77657 MW; 3660298061DB1B CRC64;

Query Match 57.4%; Score 35; DB 1; Length 708;
Best Local Similarity 54.5%; Pred. No. 89;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 PLSSTFSRIGD 11
DB 662 PLSQMFQVGD 672

RESULT 33
ID RRP1_BTV10 STANDARD; PRT; 1302 AA.
AC P13840;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE RNA-directed RNA polymerase (EC 2.7.7.48) (VP1).
GN S1.
OS Bluetongue virus (serotype 10 / isolate USA).
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OC NCBI_TaxID=10900;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89098323; PubMed=2850542;
RA Roy P., Fukusho A., Ritzer G.D., Lyon D.;
RT "Evidence for genetic relationship between RNA and DNA viruses from

```

RT the sequence homology of a putative polymerase gene of bluetongue  
RT virus with that of vaccinia virus: conservation of RNA polymerase  
RT genes from diverse species";  
RL Nucleic Acids Res. 16:11759-11767(1988).  
RN (2)  
RN SEQUENCE FROM N.A.  
RX MEDLINE=90345726; PubMed=2166648;  
RA Roy P., Marshall J.J.A., French T.J.;  
RT "Structure of the bluetongue virus genome and its encoded proteins.",  
RL Curr. Top. Microbiol. Immunol. 162:43-87(1990).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA} (N).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: X12819; CAA31306.1; -.  
CC PIR: A34296; RRRABT.  
CC InterPro: IPR007095; RNA\_pol\_DS\_PS.  
CC InterPro: IPR007097; RNA\_pol\_rec.  
CC Transference: Nucleotidyltransferase: RNA-directed RNA polymerase.  
SQ SEQUENCE 1302 AA; 149586 MW; 0C063AE396FC33D CRC64;  
  
QY Query Match 57.4%; Score 35; DB 1; Length 1302;  
Best Local Similarity 60.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
Db 539 PLSIFSRVWG 548  
1 PLSIFSRVWG 10  
||:||||:  
-----  
ID NUSG\_LIBAS STANDARD; PRT; 70 AA.  
AC P36262;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Transcription antitermination protein nusg (fragment).  
GN NUSG.  
OS Liberibacter asiaticus (Liberobacter asiaticum).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OX Rhizobiaceae; Candidatus Liberibacter.  
RN NCB1\_TaxID=34021;  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93153297; PubMed=7763375;  
RA Villedanoux S., Garnier M., Lagret F., Renaudin J., Bove J.M.;  
RT "The genome of the non-cultured, bacterial-like organism associated  
RT with citrus greening disease contains the nusg-rp1kxJL-rp0bC gene  
RT cluster and the gene for a bacteriophage type DNA polymerase.";  
RL Curr. Microbiol. 26:161-166(1992).  
CC -1- FUNCTION: INFLUENCES TRANSCRIPTION TERMINATION AND  
CC ANTITERMINATION. ACTS AS A COMPONENT OF THE TRANSCRIPTION COMPLEX,  
CC AND INTERACTS WITH THE TERMINATION FACTOR RHO AND RNA POLYMERASE  
CC (BY SIMILARITY).  
-----  
CC -1- SIMILARITY: Belongs to the nusg family.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: M94319; AAA23104.1; -.

DR InterPro: IPR005824; KOW.  
DR InterPro: IPR006646; KOW sub.  
DR InterPro: IPR001062; NusG.  
DR Pfam: PF00467; KOW; 1.  
DR SMART: SM00739; KOW; 1.  
DR PROSITE: PS01014; NUSG; 1.  
KW Transcription termination.  
FT NON TER 1  
SQ SEQUENCE 70 AA; 7749 MW; 563C62DB9A0C485 CRC64;  
  
QY Query Match 55.7%; Score 34; DB 1; Length 70;  
Best Local Similarity 45.5%; Pred. No. 12;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
  
Db 10 PVSVPFEVGE 20  
1 PVSVPFEVGE 11  
|:||||:|:  
-----  
ID NTCA\_SYNY3 STANDARD; PRT; 225 AA.  
AC P33779;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Global nitrogen regulator.  
GN NTCA OR SL1423.  
OS Synecocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
OX NCB1\_TaxID=1148;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93374671; PubMed=8366058;  
RA Frias J.B., Merida A., Herrero A., Martin-Nieto J.M., Flores E.;  
RT "General distribution of the nitrogen control gene ntca in  
RT cyanobacteria.";  
RL J. Bacteriol. 175:5710-5713(1993).  
RN (2)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hiroseawa M., Sugita M., Sasaoka S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nartuo K.,  
RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,  
RA Yamada M., Yasuda M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
CC -1- FUNCTION: REQUIRED FOR FULL EXPRESSION OF PROTEINS SUBJECT TO  
CC AMONIUM REPRESSION. TRANSCRIPTIONAL ACTIVATOR OF GENES SUBJECT TO  
CC NITROGEN CONTROL.  
CC -1- SIMILARITY: BELONGS TO THE CRP/FNR FAMILY OF TRANSCRIPTIONAL  
CC REGULATORS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: X71607; CAA50607.1; -.  
CC EMBL: D90911; BAA18011.1; -.  
CC PIR: A48640; A48640.  
CC InterPro: IPR000595; cNMP binding.  
CC InterPro: IPR001808; HTH Crp.  
CC Pfam: PF00027; cNMP\_binding; 1.  
CC Pfam: PF00325; Crp; 1.  
CC PRINTS: PR00034; HTHCRP.  
CC SMART: SM00100; cNMP; 1.



DR SMART; SM00419; HTH CRP; 1.  
 DR PROSITE; PS00042; HTH CRP FAMILY; 1.  
 DR PROSITE; PS00042; CNAME BINDING 3; 1.  
 KW Transcription regulation; DNA-Binding; Activator; Complete proteome.  
 FT NP BIND 9 131  
 FT DNA\_BIND 177 196 H-T-H MOTIF (BY SIMILARITY).  
 SQ SEQUENCE 225 AA; 25046 MW; F1EB6555757E7379 CRC64;

Query Match  
 Best Local Similarity 55.7%; Score 34; DB 1; Length 225;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LSSIFSRIG 10  
 Db 10 PLAAVFRRLG 19

RESULT 36  
 YGFF\_ECOLI STANDARD; PRT; 247 AA.

ID YGFF\_ECOLI STANDARD; PRT; 247 AA.  
 AC P52037; P71138; Q46830;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical oxidoreductase ygfF (EC 1.-.-.-).  
 GN YGFF OR B2902.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blatner F.R., Plumet G., III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12";  
 RL Science 277:1453-1474 (1997).

[2]  
 [2] SEQUENCE OF 1-74 FROM N.A.  
 RC STRAIN=K12 / W3110;  
 RX MEDLINE=94162733; PubMed=7764507;  
 RA Yamada M., Yanai S., Talkner A.;  
 RT "Analysis of products of the Escherichia coli genomic genes and  
 RT regulation of their expressions: an applicable procedure for genomic  
 RT analysis of other microorganisms";  
 RL Biosci. Biotechnol. Biochem. 58:117-120 (1994).  
 RN [3]  
 RP SEQUENCE OF 137-197 FROM N.A.

RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,  
 RA Duncan M., Federpiel N., Hyman R., Kaiman S., Komp C., Kurdi O.,  
 RA Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;  
 RL Submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases  
 CC (SDR) family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U28377; AAA69070.1; -  
 CC EMBL; U28375; AAA63083.1; -  
 CC EMBL; AE000373; AAC75940.1; -  
 CC EMBL; D21144; BA004680.1; -  
 CC EMBL; U83189; AAB040290.1; -  
 CC PIR; F65074; F65074.  
 CC HSP; P19992; IHDC.

DR Ecogene; EG12971; ygfF.  
 DR InterPro; IPR002198; ADH short.  
 DR Pfam; PF0106; adh short; 1.  
 DR PRINTS; PR00080; SDRFAMILY.  
 DR PROSITE; PS00061; ADH\_SHORT; 1.  
 KW Hypothetical protein; Oxidoreductase; Complete proteome.  
 FT NP BIND 4 28  
 FT ACT\_SITE 156 156 NAD OR NADP (BY SIMILARITY).  
 SQ SEQUENCE 247 AA; 25938 MW; 8F17257B5905E4F9 CRC64;

Query Match  
 Best Local Similarity 55.7%; Score 34; DB 1; Length 247;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LSSIFSRIGDP 12  
 Db 140 VSSVASRLGSP 150

RESULT 37  
 VGLL\_HSV6Z STANDARD; PRT; 250 AA.

ID VGLL\_HSV6Z STANDARD; PRT; 250 AA.  
 AC P52526;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Glycoprotein L precursor.  
 GN GL OR U82 OR CB10L.  
 OS Human herpesvirus (type 6 / strain 229) (HHV6).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Betaherpesvirinae; Roseolovirus.  
 OX NCBI\_TaxID=36351;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=96195263; PubMed=8634027;  
 RX Lindquester G.J., Inoue N., Allen R.D., Castelli J.W.,  
 RA Stamey F.R., Dambaugh T.R., O'Brian J.J., Danovich R.M.,  
 RA Frenkel N., Pellett P.E.;  
 RT "Restriction endonuclease mapping and molecular cloning of the human  
 RT herpesvirus 6 variant B strain 229 genome";  
 RL Arch. Virol. 141:367-379 (1996).

CC -1- FUNCTION: ASSOCIATED WITH GLYCOPROTEIN H (GH) TO FORM A COMPLEX  
 CC IMPORTANT FOR INFECTION AND CELL FUSION.  
 CC -1- SIMILARITY: TO OTHER HERPESVIRUSES GLYCOPROTEIN L.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF157706; AAB06366.1; -  
 CC PIR; T44227; T44227.  
 DR InterPro: IPR002689; Cytochrome\_glu.  
 DR Pfam; PF01801; Cytochrome\_glu; 1.  
 KW Glycoprotein; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 250  
 FT CARBOHYD 41 41 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CARBOHYD 166 166 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 250 AA; 28903 MW; 0DE7723F5019A104 CRC64;

Query Match  
 Best Local Similarity 55.7%; Score 34; DB 1; Length 250;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LSSIFSRIGDP 12  
 Db 101 LRTLFALIGDP 111

```

RESULT 38
NM_ HUMAN
ID NM_013145; STANDARD; PRT; 260 AA.
AC 013145;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Putative transmembrane protein NMA precursor.
GN NMA.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96184146; PubMed=8621228;
RA Degen W.G.U., Weternan M.A., van Groningen J.J.M.,
RA Cornelissen I.M.A.H., Lemmers J.P.W.M., Agterbos M.A.,
RA van Kessel A.G., Swart G.W.M., Bloemers H.P.J.;
RT "Expression of nma, a novel gene, inversely correlates with the
RT metastatic potential of human melanoma cell lines and xenografts.";
RL Int. J. Cancer 65:460-465(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Pearce A.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RP [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner F.S., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko U., Marusik K., Farmer A.A., Rubin G.W., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Piarne C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
RA Richarde S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Keteeman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maier M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: HIGH EXPRESSION IN KIDNEY MEDULLA, PLACENTA
CC AND SPLEEN; LOW IN KIDNEY CORTEX, LIVER, PROSTATE AND GUT. NOT
CC EXPRESSED IN NORMAL SKIN. EXPRESSION IS HIGH IN MELANOCYTES AND IN
CC 3 OUT OF 11 MELANOMA METASTASES TESTED.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION IN POORLY METASTATIC HUMAN
CC MELANOMA CELL LINES; NO EXPRESSION IN HIGHLY METASTATIC HUMAN
CC MELANOMA CELL LINES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U23070; AAC50435.1; -.
CC EMBL: AL161936; CAC16002.1; -.
CC EMBL: BC019252; AAH19252.1; -.
CC MIM: 604444; -.
CC GO: GO:0016021; C: integral to membrane; TAS.
CC GO: GO:0007048; P: oncogenesis; TAS.

```

```

KW Transmembrane; Signal; Glycoprotein.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 260 PUTATIVE TRANSMEMBRANE PROTEIN NMA.
FT TRANSMEM 153 173 POTENTIAL.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 260 AA; 29107 MW; 70C2326883CE8AB CRC64;
Query Match
Best Local Similarity 55.7%; Score 34; DB 1; Length 260;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 LSISFRIQDP 12
DB 48 LSACFSRLDP 58
RESULT 39
CB23 TOBAC
ID CB23 TOBAC STANDARD; PRT; 265 AA.
AC P27494;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chlorophyll A-B binding protein 36, chloroplast precursor (LHCII type
DE I CAB-36) (LHCP).
DE CAB36.
GN Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Havana.
RA Kavanagh T.A., Beyer M.W.;
RL Submitted (MAR-1991) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: THE LIGHT-HARVESTING COMPLEX (LHC) FUNCTIONS AS A LIGHT
RECEPTOR. IT CAPTURES & DELIVERS EXCITATION ENERGY TO PHOTOSYSTEMS
WITH WHICH IT IS CLOSELY ASSOCIATED. THE N-TERMINUS OF THE PROTEIN
EXTENDS INTO THE STROMA WHERE IT IS INVOLVED WITH ADHESION OF
GRAVAL MEMBRANES AND PHOTOREGULATED BY REVERSIBLE PHOSPHORYLATION
OF ITS THREONINE RESIDUES; BOTH ARE BELIEVED TO MEDIATE THE
DISTRIBUTION OF EXCITATION ENERGY BETWEEN PHOTOSYSTEMS I AND II.
-1- SUBUNIT: THE LHC COMPLEX CONSISTS OF CHLOROPHYLLS (A & B) AND
CHLOROPHYLL A-B BINDING PROTEINS.
-1- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X58230; CAA41188.1; -.
CC PIR: S21827; S21827.
DR InterPro: IPR001344; Chloro_Abbind.
DR Pfam: PF00504; Chloro_a-b-bind; 1.
DR ProDom: PD000275; Chloro_Abbind; 1.
DR Chlorophyll; Photosynthesis; Photosystem I; Photosystem II;
KW Thylakoid; Membrane; Chloroplast; Transil peptide; Multigene family;
KW Transmembrane; Phosphorylation.
FT TRANSIT 1 36 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 37 265 CHLOROPHYLL A-B BINDING PROTEIN 36.
SQ SEQUENCE 265 AA; 28638 MW; C9A2B59AE4F5A8E4 CRC64;
Query Match
Best Local Similarity 33.3%; Score 34; DB 1; Length 265;
Matches 4; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 1 PLUSIFSRIGDP 12

```

Db 238 PIENLFDHVADP 249

## RESULT 40

```

CE23_HORVU          STANDARD;      PRT;      268 AA.
ID CB23_HORVU
AC P27523;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chlorophyll A-B binding protein of LHClI type III, chloroplast
DE precursor (CAB).
GN LHCB.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Pallas / P-01;
RX MEDLINE=92329730; PubMed=1627782;
RA Brandt J., Nielsen V.S., Thordal-Christensen H., Simpson D.J.,
RA Okels J.S.;
RT "A barley cDNA clone encoding a type III chlorophyll a/b-binding
RT polypeptide of the light-harvesting complex II.";
RL Plant Mol. Biol. 19:699-703(1992).
CC -!- FUNCTION: THE LIGHT-HARVESTING COMPLEX (LHC) FUNCTIONS AS A LIGHT
CC RECEPTOR, IT CAPTURES & DELIVERS EXCITATION ENERGY TO PHOTOSYSTEMS
CC WITH WHICH IT IS CLOSELY ASSOCIATED. THE N-TERMINUS OF THE PROTEIN
CC EXTENDS INTO THE STROMA WHERE IT IS INVOLVED WITH ADHESION OF
CC GRANAL MEMBRANES AND PHOTOREGULATED BY REVERSIBLE PHOSPHORYLATION
CC OF ITS THREONINE RESIDUES; BOTH ARE BELIEVED TO MEDIATE THE
CC DISTRIBUTION OF EXCITATION ENERGY BETWEEN PHOTOSYSTEMS I AND II.
CC -!- SUBUNIT: THE LHC COMPLEX CONSISTS OF CHLOROPHYLLS (A & B) AND
CC CHLOROPHYLL A-B BINDING PROTEINS.
CC -!- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63197; CAA44881.1; -.
DR PIR; S22482; CDBH3.
DR InterPro; IPR001344; Chloro_Abbind.
DR Pfam; PF00504; chloro_a_b_bind; 1.
DR ProDom; PD000275; Chloro_Abbind; 1.
KW Chlorophyll; Photosynthesis; Photosystem II; Thylakoid; Membrane;
KW Chloroplast; Transit peptide; Multigene family; Transmembrane;
KW Phosphorylation.
FT TRANSIT 1 45 CHLOROPLAST (POTENTIAL).
FT CHAIN 46 268 CHLOROPHYLL A-B BINDING PROTEIN OF LHClI
FT TYPE III.
FT TRANSMEM 100 119 POTENTIAL.
FT TRANSMEM 153 173 POTENTIAL.
FT TRANSMEM 222 238 POTENTIAL.
SQ SEQUENCE 268 AA; 28759 MW; 5E48C106FC5BCAFD CRC64;

Query Match 55.7%; Score 34; DB 1; Length 268;
Best Local Similarity 41.7%; Pred. No. 50;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 PLSSIFSRIGDP 12
DB 241 PIENLFDHDDP 252

```

Search completed: November 6, 2003, 15:11:54  
Job time : 12 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using SW model

Run on: November 6, 2003, 15:09:04 ; Search time 34 Seconds  
(without alignments)  
91.077 Million cell updates/sec

Title: US-09-830-981-2  
Perfect score: 61  
Sequence: 1 PLSISFRIQDP 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

SPTREMBL 23.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Score | Query Match | Length | DB ID     | Description        |
|------------|-------|-------------|--------|-----------|--------------------|
| 1          | 61    | 100.0       | 57     | 12 Q91IC1 | Q91IC1 hepatitis b |
| 2          | 61    | 100.0       | 163    | 12 Q9WP19 | Q9WP19 hepatitis b |
| 3          | 61    | 100.0       | 281    | 12 Q41750 | Q41750 hepatitis b |
| 4          | 61    | 100.0       | 281    | 12 Q8QXP8 | Q8QXP8 hepatitis b |
| 5          | 61    | 100.0       | 378    | 12 Q67954 | Q67954 hepatitis b |
| 6          | 61    | 100.0       | 378    | 12 Q67954 | Q67954 hepatitis b |
| 7          | 61    | 100.0       | 383    | 12 Q96837 | Q96837 hepatitis b |
| 8          | 61    | 100.0       | 383    | 12 Q96840 | Q96840 hepatitis b |
| 9          | 61    | 100.0       | 389    | 12 Q8QXP9 | Q8QXP9 hepatitis b |
| 10         | 61    | 100.0       | 389    | 12 Q67875 | Q67875 hepatitis b |
| 11         | 61    | 100.0       | 389    | 12 Q9WMX3 | Q9WMX3 hepatitis b |
| 12         | 61    | 100.0       | 389    | 12 Q91CS5 | Q91CS5 hepatitis b |
| 13         | 61    | 100.0       | 389    | 12 Q90772 | Q90772 hepatitis b |
| 14         | 61    | 100.0       | 389    | 12 Q91CS2 | Q91CS2 hepatitis b |
| 15         | 59    | 96.7        | 49     | 12 Q91IC3 | Q91IC3 hepatitis b |
| 16         | 59    | 96.7        | 55     | 12 Q91XF4 | Q91XF4 hepatitis b |

|    |    |      |    |           |                    |
|----|----|------|----|-----------|--------------------|
| 17 | 59 | 96.7 | 55 | 12 Q91X62 | Q91X62 hepatitis b |
| 18 | 59 | 96.7 | 55 | 12 Q91XD1 | Q91XD1 hepatitis b |
| 19 | 59 | 96.7 | 55 | 12 Q91X89 | Q91X89 hepatitis b |
| 20 | 59 | 96.7 | 55 | 12 Q91XD0 | Q91XD0 hepatitis b |
| 21 | 59 | 96.7 | 55 | 12 Q91XA3 | Q91XA3 hepatitis b |
| 22 | 59 | 96.7 | 55 | 12 Q91X64 | Q91X64 hepatitis b |
| 23 | 59 | 96.7 | 55 | 12 Q91X83 | Q91X83 hepatitis b |
| 24 | 59 | 96.7 | 55 | 12 Q91XA1 | Q91XA1 hepatitis b |
| 25 | 59 | 96.7 | 55 | 12 Q91X86 | Q91X86 hepatitis b |
| 26 | 59 | 96.7 | 55 | 12 Q91X92 | Q91X92 hepatitis b |
| 27 | 59 | 96.7 | 55 | 12 Q91X76 | Q91X76 hepatitis b |
| 28 | 59 | 96.7 | 55 | 12 Q91XA0 | Q91XA0 hepatitis b |
| 29 | 59 | 96.7 | 55 | 12 Q91XC2 | Q91XC2 hepatitis b |
| 30 | 59 | 96.7 | 55 | 12 Q91XE9 | Q91XE9 hepatitis b |
| 31 | 59 | 96.7 | 55 | 12 Q91XB0 | Q91XB0 hepatitis b |
| 32 | 59 | 96.7 | 55 | 12 Q91XB3 | Q91XB3 hepatitis b |
| 33 | 59 | 96.7 | 55 | 12 Q91XB6 | Q91XB6 hepatitis b |
| 34 | 59 | 96.7 | 55 | 12 Q91X98 | Q91X98 hepatitis b |
| 35 | 59 | 96.7 | 55 | 12 Q91X65 | Q91X65 hepatitis b |
| 36 | 59 | 96.7 | 55 | 12 Q91XF2 | Q91XF2 hepatitis b |
| 37 | 59 | 96.7 | 55 | 12 Q91XB5 | Q91XB5 hepatitis b |
| 38 | 59 | 96.7 | 55 | 12 Q91XD9 | Q91XD9 hepatitis b |
| 39 | 59 | 96.7 | 55 | 12 Q91XA9 | Q91XA9 hepatitis b |
| 40 | 59 | 96.7 | 55 | 12 Q91X74 | Q91X74 hepatitis b |
| 41 | 59 | 96.7 | 55 | 12 Q91XD4 | Q91XD4 hepatitis b |
| 42 | 59 | 96.7 | 55 | 12 Q91X94 | Q91X94 hepatitis b |
| 43 | 59 | 96.7 | 55 | 12 Q91XC7 | Q91XC7 hepatitis b |
| 44 | 59 | 96.7 | 55 | 12 Q91X88 | Q91X88 hepatitis b |
| 45 | 59 | 96.7 | 55 | 12 Q91XC5 | Q91XC5 hepatitis b |

## ALIGNMENTS

RESULT 1  
ID Q91IC1 PRELIMINARY; PRT; 57 AA.  
AC Q91IC1;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Pre-S2 protein (Major surface antigen) (Fragment).  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_Taxid=10407;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=N314-1;  
RA Huangfu J., Deng J., Deng H.;  
RT "Preliminary Study on Pres2 Region of Hepatitis B Virus";  
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF393223; AAK84362.1; -;  
DR InterPro; IPR00349; Hepvir\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
KW NON\_TER.  
FT NON\_TER.  
SQ SEQUENCE 57 AA; 6124 MW; 8D56AF65FF849CF CRC64;  
Query Match 100.0%; Score 61; DB 12; Length 57;  
Best Local Similarity 100.0%; Pred. No. 0.00036;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 PLSISFRIQDP 12  
DB 41 PLSISFRIQDP 52  
RESULT 2  
ID Q9WP19 PRELIMINARY; PRT; 163 AA.  
AC Q9WP19;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)  
DE Large S protein (Major surface antigen) (Fragment).  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10407;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=MF;  
RA Kidd-Ljunggren K.;  
RT "Nosocomial Transmission of Hepatitis B Virus Through Multiple Dose  
RT Vials.";  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF136461; AAD44711.1; -  
DR InterPro; IPR000349; Hepvir\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
SQ SEQUENCE 163 AA; 17331 MW; 38020BE460F0BD24 CRC64;  
QY  
DB 149 PLSISFSTRIGDP 160

Query Match 100.0%; Score 61; DB 12; Length 163;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 3  
ID 041750 PRELIMINARY; PRT; 281 AA.  
AC 041750;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)  
DE Middle S protein (Major surface antigen).  
GN PRES2/S.  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10407;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=E08;  
RX MEDLINE=97368435; PubMed=9225049;  
RA Bower S., van Staden L., Kew M.C., Sim J.G.;  
RT "A unique segment of the hepatitis B virus group A genotype identified  
RT in isolates from South Africa.";  
RL EMBL; U87737; AAC58018.1; -  
DR J. Gen. Virol. 78:1719-1729(1997).  
DR InterPro; IPR000349; Hepvir\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
SQ SEQUENCE 281 AA; 31153 MW; 85D90E7B5F2C9289 CRC64;

QY  
DB 41 PLSISFSTRIGDP 52

Query Match 100.0%; Score 61; DB 12; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.002;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 4  
ID 080XP8 PRELIMINARY; PRT; 281 AA.  
AC 080XP8;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)  
DE Middle surface antigen (Major surface antigen).  
GN PRES2.  
RT Hepatitis B virus.

OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10407;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA Kay A.C.;  
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
RN (2)  
RP SEQUENCE FROM N.A.  
RA Jeanlet D., Chemin I., Mandrand B., Zoulim F., Trepo C., Kay A.;  
RT "Characterization of two Hepatitis B virus populations isolated from a  
RT Hepatitis B surface antigen-negative patient.";  
RL Hepatology 0:0-0(0).  
DR EMBL; AJ344117; CAC87014.1; -  
DR InterPro; IPR000349; Hepvir\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
SQ SEQUENCE 281 AA; 31266 MW; 37234BBD949B3B03 CRC64;  
QY  
DB 41 PLSISFSTRIGDP 52

Query Match 100.0%; Score 61; DB 12; Length 281;  
Best Local Similarity 100.0%; Pred. No. 0.002;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5  
ID 067944 PRELIMINARY; PRT; 378 AA.  
AC 067944;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)  
DE Surface proteins (Major surface antigen).  
GN PRES1/PRES2.  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10407;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA Lai M.E., Mazzoleni A.P., Portu A., Balestrieri A.;  
RT Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X85254; CAA59510.1; -  
DR InterPro; IPR000349; Hepvir\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
SQ SEQUENCE 378 AA; 41257 MW; D4370B06FFB625D5 CRC64;

QY  
DB 149 PLSISFSTRIGDP 160

Query Match 100.0%; Score 61; DB 12; Length 378;  
Best Local Similarity 100.0%; Pred. No. 0.0027;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6  
ID 067954 PRELIMINARY; PRT; 378 AA.  
AC 067954;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, last annotation update)  
DE HBV surface proteins (Major surface antigen) (Fragment).  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10407;  
RN (1)  
RP SEQUENCE FROM N.A.  
RA Lai M.E., Mazzoleni A.P., Balestrieri A.;  
RT "Sequence analysis of HBV genomes isolated from patients with HBsAg

```
RT Chronic liver disease."
RL Submitted (JAN-1994) to the EMBL/Genbank/DBJ databases.
DR EMBL; X77309; CAA54515.1; -.
DR InterPro; IPR000349; Hepvir_surfAg.
KM Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 378 AA; 41356 MW; FCI8839606E121CD CRC64;

Query Match 100.0%; Score 61; DB 12; Length 378;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 149 PLSIFSRIGDP 160

RESULT 7
096837 PRELIMINARY; PRT; 383 AA.
ID 096837
AC 096837
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Large S protein (Major surface antigen) (Fragment).
GN PRE-S/S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BH;
RA MEDLINE=97049068; PubMed=8893798;
RT "Sequence analysis of pre-S/S gene of hepatitis B virus strains of
RT genotypes A, D, and F isolated in Brazil.";
RL Arch. Virol. 141:1767-1773(1996).
DR EMBL; U55224; AAB47470.1; -.
DR InterPro; IPR000349; Hepvir_surfAg.
KM Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 383 AA; 41942 MW; 6E10F3B7471B6B1 CRC64;

Query Match 100.0%; Score 61; DB 12; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 149 PLSIFSRIGDP 160

RESULT 8
096840 PRELIMINARY; PRT; 383 AA.
ID 096840
AC 096840;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Large S protein (Major surface antigen) (Fragment).
GN PRE-S/S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BKJ;
RA MEDLINE=97049068; PubMed=8893798;
RT "Sequence analysis of pre-S/S gene of hepatitis B virus strains of
RT genotypes A, D, and F isolated in Brazil.";
DR EMBL; X65258; CAA46353.1; -.
DR InterPro; IPR000349; Hepvir_surfAg.
```

```
RL Arch. Virol. 141:1767-1773(1996).
DR EMBL; U55227; AAB47473.1; -.
DR InterPro; IPR000349; Hepvir_surfAg.
KM Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 383 AA; 42018 MW; 341AB53B7911BBDC CRC64;

Query Match 100.0%; Score 61; DB 12; Length 383;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 149 PLSIFSRIGDP 160

RESULT 9
080XP9 PRELIMINARY; PRT; 389 AA.
ID 080XP9
AC 080XP9;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Large surface antigen (Major surface antigen).
GN PRE-S1.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Kay A.C.;
RT Submitted (AUG-2001) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Jeanet D.; Chemin I.; Mandrand B.; Zoulim F.; Trepo C.; Kay A.;
RT "Characterization of two Hepatitis B virus populations isolated from a
RT Hepatitis B surface antigen-negative patient.";
RL Hepatology 0:0-0(0).
DR EMBL; AJ344117; CAC87013.1; -.
DR InterPro; IPR000349; Hepvir_surfAg.
KM Pfam; PF00695; VMSA; 1.
KW Antigen.
SQ SEQUENCE 389 AA; 42764 MW; 9AC5A4D46B73632 CRC64;

Query Match 100.0%; Score 61; DB 12; Length 389;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 149 PLSIFSRIGDP 160

RESULT 10
067875 PRELIMINARY; PRT; 389 AA.
ID 067875
AC 067875;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Pre S1/S ORF (Major surface antigen).
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RA Lai M.E.; Mazzoleni A.P.; Balestrieri A.; Mellis A.; Porru A.;
RT "Sequence analysis of HBV genomes isolated from patients with HBsAg
RT negative chronic liver disease.";
RT Submitted (MAR-1992) to the EMBL/Genbank/DBJ databases.
DR EMBL; X65258; CAA46353.1; -.
DR InterPro; IPR000349; Hepvir_surfAg.
```

DR Pfam: PF00695; VMSA; 1.  
KM Antigen.  
SQ SEQUENCE 389 AA; 42661 MW; A39542B415E46F24 CRC64;  
Query Match 100.0%; Score 61; DB 12; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 PLSSTFSRIGDP 12  
DB 149 PLSSTFSRIGDP 160  
RESULT 11  
Q9WMX3 PRELIMINARY; PRT; 389 AA.  
ID Q9WMX3  
AC Q9WMX3;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Major surface antigen.  
GN S.  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxId=10407;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1;  
RA MEDLINE=99238050; PubMed=10223539;  
RA Petrolid D.R.; Tautz B.; Wolf F.; Drescher J.;  
RT "Infection chains and evolution rates of Hepatitis B Virus in cardiac  
transplant recipients infected nosocomially.";  
RL J. Med. Virol. 58:1-10 (1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1;  
RA Petrolid D.R.;  
RT Submitted (JAN-1999) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AJ131956; CA10540.1; -  
DR InterPro; IPR000349; Hepvir\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KM Antigen.  
FT CHAIN 164 389 SURFACE ANTIGEN.  
SQ SEQUENCE 389 AA; 42708 MW; 6AA08B85DFE5B99C CRC64;  
Query Match 100.0%; Score 61; DB 12; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 PLSSTFSRIGDP 12  
DB 149 PLSSTFSRIGDP 160  
RESULT 12  
Q9IC55 PRELIMINARY; PRT; 389 AA.  
ID Q9IC55  
AC Q9IC55;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Large/middle/small S proteins (Major surface antigen).  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxId=10407;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=41;  
RA MEDLINE=21361467; PubMed=11468728;  
RA Owiredu W.K.; Kramvis A.; Kew M.C.;  
RT "Hepatitis B virus DNA in serum of healthy black African adults  
positive for hepatitis B surface antibody alone: possible association  
with recombination between genotypes A and D.";

RL J. Med. Virol. 64:441-454 (2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=41;  
RA Owiredu W.K.B.A.; Kramvis A.; Kew M.C.;  
RT Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF297619; AAK97179.1; -  
DR InterPro; IPR000349; Hepvir\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KM Antigen.  
SQ SEQUENCE 389 AA; 42580 MW; F433E04992CE6E2 CRC64;  
Query Match 100.0%; Score 61; DB 12; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 PLSSTFSRIGDP 12  
DB 149 PLSSTFSRIGDP 160  
RESULT 13  
Q90772 PRELIMINARY; PRT; 389 AA.  
ID Q90772  
AC Q90772;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Surface protein (Major surface antigen).  
GN S GENE.  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxId=10407;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Ref2;  
RA MEDLINE=98362187; PubMed=9696878;  
RA Grethe S.; Monazahian M.; Boehme I.; Thomsen R.;  
RT "Characterization of unusual escape variants of hepatitis B virus  
isolated from a hepatitis B surface antigen-negative subject.";  
RL J. Virol. 72:7692-7696 (1998).  
DR EMBL; AJ003116; CA05872.1; -  
DR InterPro; IPR000349; Hepvir\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KM Antigen.  
SQ SEQUENCE 389 AA; 42682 MW; 575E7223D068CDB CRC64;  
Query Match 100.0%; Score 61; DB 12; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 PLSSTFSRIGDP 12  
DB 149 PLSSTFSRIGDP 160  
RESULT 14  
Q9IC52 PRELIMINARY; PRT; 389 AA.  
ID Q9IC52  
AC Q9IC52;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Large/middle/small S proteins (Major surface antigen).  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxId=10407;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=57;  
RA MEDLINE=21361467; PubMed=11468728;  
RA Owiredu W.K.; Kramvis A.; Kew M.C.;  
RT "Hepatitis B virus DNA in serum of healthy black African adults



RT positive for hepatitis B surface antibody alone: possible association  
RT with recombination between genotypes A and D."  
RL J. Med. Virol. 64:441-454(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=57;  
RA Owired W.K.B.A., Krawvis A., Kew M.C.;  
RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF297620; AAK97183.1; -  
DR InterPro; IPR000349; Hepvlt\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
SQ SEQUENCE 389 AA; 42937 MW; 1709DFD7D5E01066 CRC64;  
Query Match 100.0%; Score 61; DB 12; Length 389;  
Best Local Similarity 100.0%; Pred. No. 0.0028;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 PLSISFSTRIGDP 12  
Db 149 PLSISFSTRIGDP 160

RESULT 15  
Q91IC3 PRELIMINARY; PRT; 49 AA.  
AC Q91IC3;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE Pre-S2 protein (Major surface antigen) (Fragment).  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10407;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=702-4;  
RA Huangfu J., Dong J., Deng H.;  
RL "Preliminary Study on Pres2 Region of Hepatitis B Virus.",  
Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF393221; AAK84360.1; -  
DR InterPro; IPR000349; Hepvlt\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
FT NON\_TER  
SQ SEQUENCE 49 AA; 5045 MW; 5CA3F6A356B59C3 CRC64;  
Query Match 96.7%; Score 59; DB 12; Length 49;  
Best Local Similarity 91.7%; Pred. No. 0.00075;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 PLSISFSTRIGDP 12  
Db 33 PLSISFSTRIGDP 44

RESULT 16  
Q91XF4 PRELIMINARY; PRT; 55 AA.  
AC Q91XF4;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE PreS1 protein (Major surface antigen) (Fragment).  
GN S.  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10407;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sam7;  
MEDLINE=20440648; PubMed=10970376;  
RA Flodgren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,

RA Alexeyev O.A., Kidd-Ljunggren K.;  
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:  
RT molecular analysis of prevailing hepatitis B and D virus strains."  
RL J. Clin. Microbiol. 38:3311-3316(2000).  
DR EMBL; AF247933; AAF62188.1; -  
DR InterPro; IPR000349; Hepvlt\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
FT NON\_TER  
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;  
Query Match 96.7%; Score 59; DB 12; Length 55;  
Best Local Similarity 91.7%; Pred. No. 0.00084;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 PLSISFSTRIGDP 12  
Db 41 PLSISFSTRIGDP 52

RESULT 17  
Q91X62 PRELIMINARY; PRT; 55 AA.  
AC Q91X62;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE PreS2 protein (Major surface antigen) (Fragment).  
GN S.  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10407;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sam128;  
MEDLINE=20440648; PubMed=10970376;  
RA Flodgren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,  
RA Alexeyev O.A., Kidd-Ljunggren K.;  
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:  
RT molecular analysis of prevailing hepatitis B and D virus strains."  
RL J. Clin. Microbiol. 38:3311-3316(2000).  
DR EMBL; AF247963; AAF62276.1; -  
DR InterPro; IPR000349; Hepvlt\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
FT NON\_TER  
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;  
Query Match 96.7%; Score 59; DB 12; Length 55;  
Best Local Similarity 91.7%; Pred. No. 0.00084;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 PLSISFSTRIGDP 12  
Db 41 PLSISFSTRIGDP 52

RESULT 18  
Q91XD1 PRELIMINARY; PRT; 55 AA.  
AC Q91XD1;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)  
DE PreS2 protein (Major surface antigen) (Fragment).  
GN S.  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10407;  
RN [1]  
RP SEQUENCE FROM N.A.  
STRAIN=Sam18;

```

RX MEDLINE=20440648; PubMed=10970376;
RA Flogdren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,
RA Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
RT molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247940; AAF62207.1; -
DR InterPro; IPR000349; Hepvit_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSTRIGDP 12
Db 41 PLSISFSTRIGDP 52

RESULT 19
Q91X89 PRELIMINARY; PRT; 55 AA.
AC Q91X89;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE PreS1 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_Taxid=10407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20440648; PubMed=10970376;
RA Flogdren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,
RA Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
RT molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247954; AAF62249.1; -
DR InterPro; IPR000349; Hepvit_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSTRIGDP 12
Db 41 PLSISFSTRIGDP 52

RESULT 20
Q91XDO PRELIMINARY; PRT; 55 AA.
AC Q91XDO;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE PreS1 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_Taxid=10407;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Sam33;
RX MEDLINE=20440648; PubMed=10970376;
RA Flogdren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,
RA Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
RT molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247941; AAF62212.1; -
DR InterPro; IPR000349; Hepvit_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSTRIGDP 12
Db 41 PLSISFSTRIGDP 52

RESULT 21
Q91XA3 PRELIMINARY; PRT; 55 AA.
AC Q91XA3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE PreS1 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_Taxid=10407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20440648; PubMed=10970376;
RA Flogdren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,
RA Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
RT molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247950; AAF62239.1; -
DR InterPro; IPR000349; Hepvit_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSTRIGDP 12
Db 41 PLSISFSTRIGDP 52

RESULT 22
Q91X64 PRELIMINARY; PRT; 55 AA.
AC Q91X64;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE PreS1 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.

```

OX NCBI\_TaxID=10407;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sam128;  
 RX MEDLINE=20440648; PubMed=10970376;  
 RA Floodgren E., Bengtsson S., Knutsson M., Strebekova E.A., Kidd A.H.,  
 Alexeyev O.A., Kidd-Ljunggren K.;  
 RT "Recent high incidence of fulminant hepatitis in Samara, Russia:  
 RT molecular analysis of prevailing hepatitis B and D virus strains."  
 RL J. Clin. Microbiol. 38:3311-3316(2000).  
 DR EMBL; AF247963; AAF62278.1; -  
 DR InterPro; IPR000349; Hepvit\_surfAg.  
 DR Pfam; PF00695; VMSA; 1.  
 KW Antigen.  
 FT NON\_TER  
 SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;

Query Match 96.7%; Score 59; DB 12; Length 55;  
 Best Local Similarity 91.7%; Pred. No. 0.00084;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12  
 Db 41 PLSIFSRIGDP 52

## RESULT 23

Q91X83 PRELIMINARY; PRT; 55 AA.  
 ID Q91X83;  
 AC Q91X83;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Pres2 protein (Major surface antigen) (Fragment).  
 GN S.  
 OS Hepatitis B virus.  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OX NCBI\_TaxID=10407;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sam93;  
 RX MEDLINE=20440648; PubMed=10970376;  
 RA Floodgren E., Bengtsson S., Knutsson M., Strebekova E.A., Kidd A.H.,  
 Alexeyev O.A., Kidd-Ljunggren K.;  
 RT "Recent high incidence of fulminant hepatitis in Samara, Russia:  
 RT molecular analysis of prevailing hepatitis B and D virus strains."  
 RL J. Clin. Microbiol. 38:3311-3316(2000).  
 DR EMBL; AF247963; AAF62255.1; -  
 DR InterPro; IPR000349; Hepvit\_surfAg.  
 DR Pfam; PF00695; VMSA; 1.  
 KW Antigen.  
 FT NON\_TER  
 SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;

Query Match 96.7%; Score 59; DB 12; Length 55;  
 Best Local Similarity 91.7%; Pred. No. 0.00084;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12  
 Db 41 PLSIFSRIGDP 52

## RESULT 24

Q91XAI PRELIMINARY; PRT; 55 AA.  
 ID Q91XAI;  
 AC Q91XAI;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Pres2 protein (Major surface antigen) (Fragment).  
 GN S.

OS Hepatitis B virus.  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OX NCBI\_TaxID=10407;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sam65;  
 RX MEDLINE=20440648; PubMed=10970376;  
 RA Floodgren E., Bengtsson S., Knutsson M., Strebekova E.A., Kidd A.H.,  
 Alexeyev O.A., Kidd-Ljunggren K.;  
 RT "Recent high incidence of fulminant hepatitis in Samara, Russia:  
 RT molecular analysis of prevailing hepatitis B and D virus strains."  
 RL J. Clin. Microbiol. 38:3311-3316(2000).  
 DR EMBL; AF247950; AAF62237.1; -  
 DR InterPro; IPR000349; Hepvit\_surfAg.  
 DR Pfam; PF00695; VMSA; 1.  
 KW Antigen.  
 FT NON\_TER  
 SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;

Query Match 96.7%; Score 59; DB 12; Length 55;  
 Best Local Similarity 91.7%; Pred. No. 0.00084;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12  
 Db 41 PLSIFSRIGDP 52

## RESULT 25

Q91X86 PRELIMINARY; PRT; 55 AA.  
 ID Q91X86;  
 AC Q91X86;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Pres2 protein (Major surface antigen) (Fragment).  
 GN S.  
 OS Hepatitis B virus.  
 OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
 OX NCBI\_TaxID=10407;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sam82;  
 RX MEDLINE=20440648; PubMed=10970376;  
 RA Floodgren E., Bengtsson S., Knutsson M., Strebekova E.A., Kidd A.H.,  
 Alexeyev O.A., Kidd-Ljunggren K.;  
 RT "Recent high incidence of fulminant hepatitis in Samara, Russia:  
 RT molecular analysis of prevailing hepatitis B and D virus strains."  
 RL J. Clin. Microbiol. 38:3311-3316(2000).  
 DR EMBL; AF247955; AAF62252.1; -  
 DR InterPro; IPR000349; Hepvit\_surfAg.  
 DR Pfam; PF00695; VMSA; 1.  
 KW Antigen.  
 FT NON\_TER  
 SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;

Query Match 96.7%; Score 59; DB 12; Length 55;  
 Best Local Similarity 91.7%; Pred. No. 0.00084;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12  
 Db 41 PLSIFSRIGDP 52

## RESULT 26

Q91X92 PRELIMINARY; PRT; 55 AA.  
 ID Q91X92;  
 AC Q91X92;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Pres2 protein (Major surface antigen) (Fragment).  
 GN S.

GN S.  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10407;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sam76;  
RX MEDLINE=20440648; PubMed=10970376;  
RA Flogdren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,  
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:  
RL molecular analysis of prevailing hepatitis B and D virus strains.";  
J. Clin. Microbiol. 38:3311-3316(2000).  
DR EMBL; AF247953; AAF62246.1; -  
DR InterPro; IPR000349; Hepvit\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
FT NON\_TER  
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;  
Query Match  
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;  
Pred. No. 0.00084;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLUSIFSRIQDP 12  
Db 41 PISSIFSRIQDP 52

## RESULT 27

Q9IX76 PRELIMINARY; PRT; 55 AA.  
AC Q9IX76;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE Pres1 protein (Major surface antigen) (Fragment).  
GN S.  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10407;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sam104;  
RX MEDLINE=20440648; PubMed=10970376;  
RA Flogdren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,  
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:  
RL molecular analysis of prevailing hepatitis B and D virus strains.";  
J. Clin. Microbiol. 38:3311-3316(2000).  
DR EMBL; AF247959; AAF62266.1; -  
DR InterPro; IPR000349; Hepvit\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
FT NON\_TER  
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;

Query Match  
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;  
Pred. No. 0.00084;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLUSIFSRIQDP 12  
Db 41 PISSIFSRIQDP 52

## RESULT 28

Q9IXA0 PRELIMINARY; PRT; 55 AA.  
AC Q9IXA0;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE Pres1 protein (Major surface antigen) (Fragment).  
GN S.  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10407;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sam68;  
RX MEDLINE=20440648; PubMed=10970376;  
RA Flogdren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,  
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:  
RL molecular analysis of prevailing hepatitis B and D virus strains.";  
J. Clin. Microbiol. 38:3311-3316(2000).  
DR EMBL; AF247951; AAF62242.1; -  
DR InterPro; IPR000349; Hepvit\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
FT NON\_TER  
SQ SEQUENCE 55 AA; 5892 MW; 4AB600FC488BFF73 CRC64;  
Query Match  
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;  
Pred. No. 0.00084;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLUSIFSRIQDP 12  
Db 41 PISSIFSRIQDP 52

## RESULT 29

Q9IXC2 PRELIMINARY; PRT; 55 AA.  
AC Q9IXC2;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
DE Pres2 protein (Major surface antigen) (Fragment).  
GN S.  
OS Hepatitis B virus.  
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.  
OX NCBI\_TaxID=10407;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Sam38;  
RX MEDLINE=20440648; PubMed=10970376;  
RA Flogdren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,  
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:  
RL molecular analysis of prevailing hepatitis B and D virus strains.";  
J. Clin. Microbiol. 38:3311-3316(2000).  
DR EMBL; AF247943; AAF62216.1; -  
DR InterPro; IPR000349; Hepvit\_surfAg.  
DR Pfam; PF00695; VMSA; 1.  
KW Antigen.  
FT NON\_TER  
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;

Query Match  
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;  
Pred. No. 0.00084;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLUSIFSRIQDP 12  
Db 41 PISSIFSRIQDP 52

## RESULT 30

Q9IXE9 PRELIMINARY; PRT; 55 AA.  
AC Q9IXE9;

```
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Pres2 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sam9;
RX MEDLINE=20440648; PubMed=10970376;
RA Floodgren E., Bengtsson S., Knutsson M., Strechkova E.A., Kidd A.H.,
  Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
  molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247934; AAF62189.1; -
DR InterPro; IPR000349; Hepvit_surfAg.
DR Pfam; PF00695; VWSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 55 AA; 5863 MW; 4AE61FEC488BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFSTRIGDP 12
Db 41 PLSISFSTRIGDP 52

RESULT 31
Q91XB0 PRELIMINARY; PRT; 55 AA.
AC Q91XB0;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Pres2 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sam58;
RX MEDLINE=20440648; PubMed=10970376;
RA Floodgren E., Bengtsson S., Knutsson M., Strechkova E.A., Kidd A.H.,
  Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
  molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247947; AAF62228.1; -
DR InterPro; IPR000349; Hepvit_surfAg.
DR Pfam; PF00695; VWSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 55 AA; 5863 MW; 4AE61FEC488BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFSTRIGDP 12
Db 41 PLSISFSTRIGDP 52

RESULT 32
Q91XB3 PRELIMINARY; PRT; 55 AA.
```

```
AC Q91XB3;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Pres2 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sam52;
RX MEDLINE=20440648; PubMed=10970376;
RA Floodgren E., Bengtsson S., Knutsson M., Strechkova E.A., Kidd A.H.,
  Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
  molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247946; AAF62225.1; -
DR InterPro; IPR000349; Hepvit_surfAg.
DR Pfam; PF00695; VWSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 55 AA; 5863 MW; 4AE61FEC488BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFSTRIGDP 12
Db 41 PLSISFSTRIGDP 52

RESULT 33
Q91XB6 PRELIMINARY; PRT; 55 AA.
AC Q91XB6;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Pres2 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sam12;
RX MEDLINE=20440648; PubMed=10970376;
RA Floodgren E., Bengtsson S., Knutsson M., Strechkova E.A., Kidd A.H.,
  Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
  molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247935; AAF62192.1; -
DR InterPro; IPR000349; Hepvit_surfAg.
DR Pfam; PF00695; VWSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 55 AA; 5863 MW; 4AE61FEC488BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSISFSTRIGDP 12
Db 41 PLSISFSTRIGDP 52

RESULT 34
Q91XB8 PRELIMINARY; PRT; 55 AA.
```

```
ID Q9IX98 PRELIMINARY; PRT; 55 AA.
AC Q9IX98;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Pres2 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sam68;
RX MEDLINE=20440648; PubMed=10970376;
RA Floegren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,
  Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
  molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247951; AAF62240.1; -
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 55 AA; 5892 MW; 4AB600FC48BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSTRIGDP 12
DB 41 PLSISFSTRIGDP 52

RESULT 35
Q9IX65 PRELIMINARY; PRT; 55 AA.
ID Q9IX65;
AC Q9IX65;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Pres2 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sam125;
RX MEDLINE=20440648; PubMed=10970376;
RA Floegren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,
  Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
  molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247962; AAF62273.1; -
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC48BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSTRIGDP 12
DB 41 PLSISFSTRIGDP 52

RESULT 36
```

```
Q9IXF2 PRELIMINARY; PRT; 55 AA.
ID Q9IXF2;
AC Q9IXF2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Pres2 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sam7;
RX MEDLINE=20440648; PubMed=10970376;
RA Floegren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,
  Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
  molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247933; AAF62186.1; -
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC48BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSTRIGDP 12
DB 41 PLSISFSTRIGDP 52

RESULT 37
Q9IXB5 PRELIMINARY; PRT; 55 AA.
ID Q9IXB5;
AC Q9IXB5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Pres1 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sam52;
RX MEDLINE=20440648; PubMed=10970376;
RA Floegren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,
  Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
  molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247946; AAF62227.1; -
DR InterPro; IPR000349; Hepvir_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC48BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSISFSTRIGDP 12
DB 41 PLSISFSTRIGDP 52
```

```

RESULT 38
Q9IXD9 PRELIMINARY; PRT; 55 AA.
AC Q9IXD9;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Pres1 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sam15;
RX MEDLINE=20440648; PubMed=10970376;
RA Flodgren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,
RA Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
RT molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247938; AAF62203.1; -.
DR InterPro; IPR000349; Hepvit_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
FT NON_TER
SQ SEQUENCE 55 AA; 5892 MW; 52514FEC488BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 41 PLSIFSRIGDP 52

RESULT 39
Q9IXA9 PRELIMINARY; PRT; 55 AA.
AC Q9IXA9;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Pres1 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sam59;
RX MEDLINE=20440648; PubMed=10970376;
RA Flodgren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,
RA Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
RT molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247948; AAF62233.1; -.
DR InterPro; IPR000349; Hepvit_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
FT NON_TER
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 41 PLSIFSRIGDP 52

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12

```

```

Db 41 PLSIFSRIGDP 52

RESULT 40
Q9IX74 PRELIMINARY; PRT; 55 AA.
AC Q9IX74;
DT 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DE 01-OCT-2002 (TREMblrel. 22, Last annotation update)
DE Pres2 protein (Major surface antigen) (Fragment).
GN S.
OS Hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sam104;
RX MEDLINE=20440648; PubMed=10970376;
RA Flodgren E., Bengtsson S., Knutsson M., Strebkova E.A., Kidd A.H.,
RA Alexeyev O.A., Kidd-Ljunggren K.;
RT "Recent high incidence of fulminant hepatitis in Samara, Russia:
RT molecular analysis of prevailing hepatitis B and D virus strains.";
RL J. Clin. Microbiol. 38:3311-3316(2000).
DR EMBL; AF247959; AAF62264.1; -.
DR InterPro; IPR000349; Hepvit_surfAg.
DR Pfam; PF00695; VMSA; 1.
KW Antigen.
FT NON_TER
FT NON_TER
SQ SEQUENCE 55 AA; 5863 MW; 4AB61FEC488BFF73 CRC64;

Query Match
Best Local Similarity 96.7%; Score 59; DB 12; Length 55;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSIFSRIGDP 12
Db 41 PLSIFSRIGDP 52

Search completed: November 6, 2003, 15:12:42
Job time : 36 secs

```

**THIS PAGE BLANK (USPTO)**